DYNAMIC®
LEARNING MAPS

# *2020–2021 DLM Administration During COVID-19: Participation, Performance, and Educational Experience*

Technical Report #21-02

**December 2021**

Accessible Teaching, Learning, and Assessment Systems. (2021). *2020–2021 DLM Administration During COVID-19: Participation, Performance, and Educational Experience* (Technical Report No. 21-02). University of Kansas.

**Acknowledgements**

Many contributors made the writing of this technical report possible. Dynamic Learning Maps® (DLM®) staff who made significant writing contributions to this report are listed below with gratitude in alphabetical order.

**Brianna Beitling, M.S.**, *Psychometrician Associate*
**Amy K. Clark, Ph.D.**, *Associate Director for Operational Research*
**Ashley Hirt, Ph.D.**, *Research Project Specialist*
**Jeffrey C. Hoover, M.S.**, *Psychometrician Associate*
**Meagan Karvonen, Ph.D.**, *Director of Accessible Teaching, Learning, and Assessment Systems*
**Jennifer Kobrin, Ed.D.**, *Research Project Director*
**Noelle Pablo**, *Psychometrician Assistant*
**W. Jake Thompson, Ph.D.**, *Senior Psychometrician*

The authors wish to acknowledge Brooke Nash for her valuable review, which helped strengthen this report. We would also like to acknowledge Tina Wang for her data analysis, as well as Elizabeth Kavitsky and Justin Dean for their editing support.

# Contents

# List of Tables

# List of Figures

# Executive Summary

COVID-19 prompted the closure of schools throughout the United States in March 2020, resulting in a significant impact on instruction, assessment, and student learning. Schools re-opened in the fall of 2020 under a variety of instructional scenarios, including remote, hybrid, and in-person models, within and across states. Instructional models changed throughout the 2020–2021 academic year, based on factors such as COVID-19 infection rates, student needs, district size, state and local policy, and parent choice. Instructional changes inevitably impacted student learning. To better understand the effects of the pandemic and the associated instructional changes, staff at Accessible Teaching, Learning, and Assessment Systems (ATLAS) conducted research on student participation, educational experiences, and assessment performance for students taking Dynamic Learning Maps® (DLM®) assessments during the 2020–2021 school year, using a variety of data sources, including teacher survey responses, student achievement data, and enrollment records. Key findings include:

- Around 70,000 students completed at least one DLM assessment; more than 110,000 might have been expected in a typical year.
- Participation varied by state, from a low of 51% to a high of 95% of the number of students who participated in 2018–2019.
- Students from historically marginalized populations and those at the lowest complexity level (i.e., those who typically are the lowest achieving) participated at lower rates than in previous years.
- Teachers reported nearly half of students taking DLM assessments spent 75%–100% of their instructional time in school, but that at least 60% of students spent some portion of the year receiving instruction from their teacher in home, and 30% of students received instruction from a family member. Only 6% of students received no in-school instruction.
- More than 70% of students experienced changes between remote and in-person learning at least once during the academic year.
- Survey results suggest that the pandemic affected students' opportunity to learn; however, these results may reflect changes in the student population tested in 2020–2021 compared to previous years.
- Student were less likely to receive depth (i.e., more than 20 hours) and breadth (i.e., more than 5 hours) of instruction in all English language arts and mathematics conceptual areas, science core ideas, and science and engineering practices.
- Though survey data suggest a decline in breadth and depth of instruction from 2018–2019 to 2020–2021, students with matched survey data demonstrated the same or greater amount of instruction in 2020–2021 compared to 2018–2019. This may suggest that teachers provided instruction for a narrower range of content in 2020–2021 (i.e., spent more time on a fewer number of topics).
- Blueprint coverage remained stable (~95%) for students in states adopting the Year-End model and for science, compared to previous years. Coverage for students in states adopting the Instructionally Embedded model increased to around 95% as well.
- The use of a propensity score model to control for population changes across years indicates that there were factors other than demographic composition that impacted student performance in 2020–2021.
- Model changes implemented in 2019–2020 make evaluating COVID-related impacts on performance challenging in English language arts and mathematics. For these subjects, the Instructionally Embedded model observed a general decrease in performance, while the Year-End model observed improvements from the previous administration. Science, which did not have any model changes, observed a decrease in performance relative to prior administrations. Taken together, these findings suggest the COVID-19 pandemic did have some impact on overall student performance, even after accounting for population shifts over time.
- Additional research is needed to determine the full impact of COVID-19 on DLM assessments and the student population across years.

# 1. Introduction

This report describes the administration of the Dynamic Learning Maps® (DLM®) alternate assessments during the 2020–2021 school year. In this report, we compare students' participation, educational experience, and performance in 2020–2021 to prior years to understand administration during the COVID-19 pandemic. However, in addition to the COVID-19 pandemic, there are many factors that may influence these comparisons across years. For example, state compliance with the Every Student Succeeds Act (ESSA, 2015) 1% threshold for alternate assessments based on alternate achievement standards [AA-AAS] participation may affect participation across years. Similarly, changes to assessment blueprint and administration practice may have an effect on student performance. Throughout the report, we discuss the confounding factors that make it difficult to evaluate effects that are specific to COVID-19. Thus, readers are cautioned against making causal inferences about the impact of COVID-19 on DLM administration.

## 1.1. COVID-19 Response

The COVID-19 pandemic had a significant impact on instruction, learning, and assessment. Beginning in March 2020, in response to the growing pandemic and recommendations by the Centers for Disease Control and Prevention (CDC, 2020), many states and local school districts began to close to slow the spread of the virus. During the school closures, students were unable to complete assessments, leading the United States Secretary of Education to invite states to submit 1-year waivers of assessment and accountability requirements of the Elementary and Secondary Education Act of 1965 (ESEA, 1965), as amended by ESSA (2015). All 50 states, the District of Columbia, the Commonwealth of Puerto Rico, and the Bureau of Indian Education applied for and received this waiver for 2019–2020 assessments (*Recommended Waiver Authority Under Section 3511(d)(4) of Division A of the Coronavirus Aid, Relief, and Economic Security Act ("CARES ACT")*, 2020).

Following the complete school and district closures and the halting of assessment administration in the spring of 2020, the reopening of schools in fall 2020 was characterized by variations of remote, in-person, and hybrid instructional models both within and across states. In many states and districts, the degree to which these instructional models were utilized changed over the course of the school year and was dependent on multiple factors including COVID-19 case counts, district size, ages of students within schools, local policy, student needs, and parent choice. While state and local education agencies made every effort to ensure all students had access to instruction and instructional materials regardless of learning environment, it is well acknowledged that changes to learning inevitably occurred during the 2020–2021 academic year. Recognizing both the variability of instructional access and state and local need for data on student achievement, on February 22, 2021, the U.S. Department of Education's Office of Elementary and Secondary Education provided states with guidance regarding assessment, accountability, and reporting requirements for the 2020–2021 school year. The department's guidance, as it relates to assessments, offered states the option to apply for a 1-year waiver from accountability requirements as well as flexibility in assessment administration. The types of flexibility described in the department's letter included administering shorter versions of state assessments, offering remote administration where feasible, and extending testing windows. The guidance further explained that the focus of this year's assessments is "to provide information to parents, educators, and the public about student performance and to help target resources and supports" (Rosenblum, 2021).

Throughout spring 2020 and into the 2020–2021 school year, staff at Accessible Teaching, Learning, and Assessment Systems (ATLAS), who facilitate the DLM Consortium, took several steps to support student learning and the validity of interpretations made from DLM assessment results. To support remote, in-person, and hybrid learning models, ATLAS released 50 additional testlets (i.e., short assessments of three to nine items) in spring 2020 to be used as instructional resources by parents, teachers, and students. ATLAS staff also established and facilitated a Remote Learning Ad Hoc committee among DLM member states in which state education agency staff could collaborate and share strategies that were working within their states. ATLAS staff also consulted with state education agency staff and the DLM Technical Advisory Committee (TAC) to evaluate how different instruction and

assessment scenarios would impact the level of reporting that could be supported on summative score reports. For a summary of this work, see Clark et al. (2021). ATLAS staff, at the recommendation of the DLM TAC, also issued a policy statement against remote test administration (i.e., in a virtual environment such as Zoom, Microsoft Teams, Google Hangouts, etc., in which the test administrator is not physically present during administration). This policy did not preclude assessments from being administered at an off-site location, such as the student's home or a testing facility. Due to the unique accessibility needs of the student population taking DLM assessments and the supports typically provided to students during assessment, the policy indicated that DLM assessments should only be administered in person by trained test administrators.

During the 2020–2021 school year, ATLAS staff continued to communicate with state education agencies to gather information on how instruction and assessment were proceeding in their states. Overall, there was a wide range of instruction and assessment scenarios both within and across states, with some school districts entirely remote for the majority of the year and others focusing on returning to in-person instruction as soon as possible. Additionally, states had varying policies regarding DLM assessment administration windows. Some states operated similarly to previous years, and others extended their spring testing window. In two states, no spring assessments were administered. This variability in instruction and assessment poses challenges for evaluating student participation, experience, and performance across years, as the specific impact of COVID-19 is highly dependent on the response in each locality.

## 1.2. Assessment Administration Changes

In addition to pandemic-related impacts on instruction and assessment in 2020–2021, the DLM Consortium also implemented changes to administration models beginning in the 2019–2020 academic year.

States administering English language arts (ELA) and mathematics assessments choose between Instructionally Embedded and Year-End administration models. All states participating in science follow a single administration model for science assessments, regardless of which model they use for ELA and mathematics.

At the start of the 2019–2020 school year, prior to the COVID-19 pandemic, several changes to the administration of DLM assessments were implemented for both the Instructionally Embedded and Year-End ELA and mathematics assessment models. Because these changes coincided with the pandemic, the administration changes act as a confounding variable in differentiating causal impacts of COVID-19 on instruction and assessment. Throughout the report, we point out which results are likely a result of disruptions due to COVID-19 and which might have other explanations, such as assessment changes. Here we provide a high-level overview of the changes that were implemented in 2019–2020 to provide additional context.

### *1.2.1. Instructionally Embedded Model Changes*

For Instructionally Embedded ELA and mathematics assessments, students complete assessments throughout the school year during fall and spring administration windows. Results from the entire year are used for summative scoring and reporting. Short assessments, called testlets, measure student knowledge relative to alternate content standards, known as Essential Elements (EEs). For each EE, assessments are available at five linkage levels that vary in complexity from the grade-level expectation defined in the EE (Initial Precursor, Distal Precursor, Proximal Precursor, Target, Successor[1]). The blueprint specifies the number of EEs for which students are expected to complete assessments from among the full set of available EEs. Teachers decide the EEs and levels on which each student is assessed.

In 2019–2020, states adopting the Instructionally Embedded model approved changes to the administration process. Prior to that year, students were expected to cover the full blueprint in the fall administration window and were assessed on a subset of EEs (four to five total) in the spring window. Beginning in 2019–2020, the administration

[1] Science administers testlets at three linkage levels (Initial, Precursor, and Target).

procedure was updated such that teachers covered the full blueprint in both administration windows (i.e., the full blueprint is covered twice), with the spring administration functioning identically to fall. Along with this administration change, a new online tool, the Instruction and Assessment Planner, was developed to aid teachers in making their assessment choices. The Planner tool more clearly indicated which blueprint requirements had and had not been met and recommended the linkage level to administer for each EE. The linkage level recommended in fall was based off responses to the student's First Contact survey (the survey is described in more detail in the Data Sources section); the spring recommendation was based off available fall response data. For a complete description of the Instructionally Embedded administration changes and the Instruction and Assessment Planner, see Chapter 4 of DLM Consortium (2020a).

### 1.2.2. Year-End Model Changes

In states adopting the Year-End assessment model, testlets are also available to assess the EEs at five linkage levels. A test blueprint specifies the full set of EEs for which students are expected to complete assessments. There is no teacher choice in administration. The assessment system assigns testlets during a spring window to cover all blueprint requirements.

When specifying the blueprint, Year-End model states originally prioritized broad content sampling while also limiting the number of assessment items to avoid response fatigue. However, this limited the number of items available to measure each EE. In 2019–2020, states participating in the Year-End model adopted changes to the assessment blueprint that reduced the total number of EEs measured by the assessment. By reducing the number of assessed EEs, students complete more items per EE, increasing the certainty of individual skill mastery determinations. As a result, individual student score reports were updated to include the Learning Profile that summarizes fine-grained mastery status for assessed EEs and linkage levels (see Chapter 7 of DLM Consortium, 2020b). The change also resulted in a change to the pool of operational content. Under the original blueprint, multiple EEs were assessed on a single testlet. With the revised blueprint, testlets measured only one EE. These single-EE testlets were already in operational use for Instructionally Embedded ELA and mathematics assessments. For a complete description of the blueprint changes, see Chapter 3 of DLM Consortium (2020b).

The changes to the blueprint and item pool necessitated an adjustment to the number of linkage levels needed to reach each achievement standard. For DLM assessments, a student's overall achievement level is determined by applying cut points to the total number of linkage levels the student mastered. Because the number of EEs on the Year-End ELA and mathematics blueprints was reduced, the total number of linkage levels possible was similarly reduced. Thus, the original cut points were no longer appropriate. In February 2020, preliminary cut points based on projected performance under the revised blueprints were shared with the DLM TAC. The intention was for states to use Learning Profiles and impact data from 2019–2020 to review and refine the preliminary cut points prior to 2020 summative reporting. Because assessments were cancelled in spring 2020, this was not possible. Additionally, due to concerns about sample representation and disrupted instruction, the DLM TAC recommended an administrative standard setting using the preliminary cut points for 2020–2021 for the purpose of reporting results that could support student learning, consistent with the U.S. Department of Education's guidance to states. The full standards adjustment process, including state review of impact data and Learning Profiles, was postponed to the 2021–2022 assessment administration.

### 1.2.3. Summary of Administration Changes

In summary, both the Instructionally Embedded and Year-End ELA and mathematics assessments had changes intended for the 2019–2020 administration that were implemented in full in 2020–2021 due to pandemic closures during spring 2020. These changes are summarized in Table 1. The Instructionally Embedded model saw changes to how testlets are administered in the spring window and the introduction of the new Planner tool to help teachers

make instruction and assessment decisions. The Year-End model adopted revised assessment blueprints, which were accompanied by changes to the operational testlet pool and cut points for determining the overall achievement level. The potential impacts of these changes are discussed throughout the report.

**Table 1**

*Summary of Assessment Changes From 2018–2019 to 2020–2021*

| Instructionally Embedded | Year-End | Science |
|---|---|---|
| • Planner tool for selecting EEs and levels for instruction and assessment (Fall 2019)<br>• Spring blueprint coverage requirements (Spring 2020)<br>• Teacher selection of spring EEs and linkage levels (Spring 2020) | • ELA and mathematics blueprint revisions (Fall 2019)<br>• Operational item pool changes from multi-EE to single-EE testlets (Spring 2020)<br>• Administrative standards adjustments using preliminary cut points (Spring 2021) | • No changes |

*Note.* Parentheses represent when the change was implemented.

## 1.3. Purpose of the Report

The purpose of this report is to use several data sources, including enrollment records, teacher survey responses, and achievement data, to summarize DLM administration during the COVID-19 pandemic (i.e., 2020–2021 academic year) for students with the most significant cognitive disabilities who take DLM assessments. We first describe the student population, including both who participated in DLM assessments and who did not. We then describe findings related to students' experiences with both instruction and assessment. Finally, we describe performance for students who took DLM alternate assessments in ELA, mathematics, and science. Research questions are organized around these three areas.

1. Who participated in DLM assessments, and how did the population compare to previous years?
2. What were students' educational experiences during the 2020–2021 school year, and how did they compare to previous years?
3. How did students perform in 2020–2021, and how did their performance compare to previous years?
    a. How were changes in performance related to other changes to the assessment administration?

# 2. Data Sources

Data were collected from system enrollment records, teacher responses to the First Contact survey of learner characteristics, the annual spring teacher survey, special circumstances files, and DLM assessments. Throughout the report, we use the full set of data available in each year. For example, some states that are currently members of the DLM Consortium were not members in 2017–2018 and/or 2018–2019. Thus, those states do not have data for the prior years. Rather than exclude those states due to the lack of data, we have included all of the data available for each administration year. Removing those states would have resulted in a consistent sample for all comparison years; however, the removal would also have limited a 2020–2021 sample that was already reduced due to low participation (described below in the Participation in DLM Assessments section). We also recognize that the COVID-19 pandemic affected states and districts in various ways. By removing states that did not participate in all years, we may have unintentionally masked or amplified these effects, depending on a number of factors, such as local policy, severity of outbreaks, and the resources available for schools to adapt to remote instruction. Additionally, our purpose in this report was to describe the participation, experience, and performance of all students who did participate, including those students from states that did not participate in all years. Future research could examine a related but distinct question about the differences in a matched sample across years. Therefore, we made the decision to include all states that participated in each administration year.

In the following sections we describe the data that are collected and used throughout the report.

## 2.1. Enrollment Records

Student demographic data are collected as part of enrollment in DLM assessments. Data include gender, race, Hispanic ethnicity, and enrollment or participation in English learner (EL) services.

## 2.2. First Contact Survey

The First Contact Survey is a survey of learner characteristics that is completed for each student before any assessment items are administered. The survey is completed by the student's teacher and includes items related to the student's knowledge, skills, and understandings in ELA, mathematics, science, and expressive communication. These responses are used to calculate a complexity band for each student. There are four bands for each subject: Foundational, Band 1, Band 2, and Band 3. Due to the heterogeneous population of students who participate in DLM assessments, the complexity band serves as a more useful variable for exploring results than disability label, which may be impacted by local implementation policies. Within the assessments, complexity bands are used to recommend linkage levels for the Instructionally Embedded ELA and mathematics assessments and are used to select the starting linkage level for the science and Year-End ELA and mathematics assessments. Previous work described by Nash and Thompson (2017) has indicated that the complexity bands are effective for matching students to appropriate linkage levels. Additional First Contact survey items collect information about learning characteristics for the student (e.g., receptive communication, attention). For a complete description of the First Contact Survey, see Nash et al. (2016).

## 2.3. Teacher Survey

Each spring, a DLM teacher survey is administered to collect information about student and teacher experiences. One survey is administered per student in the Kite Student Portal to be filled out by the teacher to whom the student is assigned to in the system. The 2020–2021 survey consisted of four blocks. Three of the blocks were provided for all students to collect information about student experience with the assessment, teacher background, and instruction and assessment conditions during the year. The other block was spiral-assigned, with four forms available. Teachers were asked questions about accessibility of the assessment or about students' opportunity to learn in one of three subjects (i.e., amount of instructional time in ELA, mathematics, or science).

In 2020–2021, a total of 14,681 teachers responded to the teacher survey (response rate of 63% of teachers) about 39,661 students' experiences (51% of students). From 2018–2019 to 2020–2021, the response rate at the student level declined by 15 percentage points, and the teacher response rate declined by 14 percentage points. Teacher responses represented fewer districts and schools as well, with declines of 6 and 11 percentage points, respectively. Thus, throughout the report, responses to the survey may not be representative of the entire population of teachers administering DLM assessments due to non-responses.

Although the teacher survey was available for all students, including students who did not take any DLM assessments, only a very small number of students who were not tested are represented in the 2020–2021 survey data (0.06%, $n$ = 24), which is a similar rate as was observed in 2017–2018 (0.06%, $n$ = 37) and 2018–2019 (0.05%, $n$ = 28).

## 2.4. Special Circumstances Files

Special circumstances files are delivered to state education agencies following the close of the spring assessment window. These files are used by state education agencies to help inform decisions about which student records, if any, should be invalidated. Special circumstance codes are applied to specific testlets by the teacher or state or local education agency staff. A student may have multiple special circumstance codes in a subject if multiple testlets have codes applied; alternately, one applied code may explain why no testlets or subsequent testlets were not completed (e.g., chronic absences). The files delivered to state education agencies provide information about which students and EEs were affected by extenuating circumstances, as defined by each state. Three codes were added for the 2020–2021 academic year: 1. Student could not test due to COVID-19; 2. Teacher administered the assessment remotely; and 3. Non-teacher administered.

## 2.5. DLM Assessments

Results from the DLM assessments are reported as fine-grained mastery decisions for each assessed EE and linkage level, which are reported in the Learning Profile. The linkage level mastery classifications are also aggregated to support inferences at larger grain sizes, including the percentage of linkage levels mastered within each claim or conceptual area, and an overall achievement level based on the total number of linkage levels mastered for each subject. The achievement level is used in state accountability systems and provides a summative description of each student's achievement in each subject. For a complete description of DLM assessment results, see Chapter 7 of DLM Consortium (2016a).

# 3. Participation in DLM Assessments

We examined DLM assessment participation over a 4-year period during which new states joined the consortium and the COVID-19 pandemic impacted statewide assessments. Table 2 shows the number of students, educators, schools, districts, and states that participated in DLM assessments in 2017–2018, 2018–2019, and 2020–2021. Counts presented for 2020–2021 are based on data collected through July 6, 2021. Two states extended the spring testing window through September 2021. However, because fall administration of spring assessments likely occurred under different circumstances (e.g., new grade, possible new teacher), we chose not to include those data in this report.

**Table 2**

*Participation Counts for 2017–2018, 2018–2019, and 2020–2021*

| Group | 2017–2018 | 2018–2019 | 2020–2021 |
|---|---|---|---|
| Students | 86,471 | 92,073 | 70,038 |
| Educators | 23,909 | 24,965 | 22,839 |
| Schools | 13,338 | 14,048 | 13,459 |
| Districts | 4,167 | 4,395 | 4,707 |
| States | 17 | 19 | 19 |

*Note*. 2019–2020 omitted due to nationwide cancellation of large-scale assessment administration.

In 2020–2021, DLM alternate assessments were administered to students in 19 states. Note that although two states did not participate in 2020–2021, there were two additional states that joined the DLM Consortium between 2018–2019 and 2020–2021. This is why we do not see a drop from 19 states to 17 in 2020–2021. Had all states participated, there would be a total of 21 states in 2020–2021. Despite the addition of those two new states, there were fewer students, educators, and schools participating in 2020–2021 than in prior years. A total of 70,038 students participated in the DLM assessment in 2020–2021. Factoring in the two additional states, over 110,000 students might have been expected in a typical year.

Table 3 shows the number of students for each state who participated in 2020–2021, as well as the percentage of the 2018–2019 and 2017–2018 sample the 2020–2021 sample represents within each state. In some states, the number of students participating in 2020–2021 was as low as 51% of the number of students in 2018–2019. In others, participation was as high as 95% of the 2018–2019 students. It should be noted that state-level participation may also be impacted ongoing efforts to comply with the ESSA (2015) 1% threshold on students participating in alternate assessments. Of 4,395 districts that had students enrolled in 2018–2019, 68% had a decrease in students from 2018–2019 to 2020–2021. The 10 largest districts accounted for 23% of the decrease, and the largest 20% of districts accounted for 80% of the decrease, suggesting urban districts were responsible for the majority of the decrease in participation in 2020–2021. Across districts with decreases, participation in 2020–2021 was typically about half of what was observed in 2018–2019.

**Table 3**

*Student Participation, by State*

| State | Total Students (*n*) | | | 2020–2021 (*n*) as % of | |
|---|---|---|---|---|---|
| | 2017–2018 | 2018–2019 | 2020–2021 | 2017–2018 | 2018–2019 |
| Alaska | 622 | 563 | 385 | 61.9 | 68.4 |
| Arkansas | * | 4,415 | 2,463 | — | 55.8 |
| Colorado | 5,224 | 5,071 | 3,100 | 59.3 | 61.1 |
| Delaware | 1,237 | 1,229 | 715 | 57.8 | 58.2 |
| Illinois | 11,524 | 14,679 | 7,513 | 65.2 | 51.2 |
| Iowa | 2,920 | 2,988 | 2,654 | 90.9 | 88.8 |
| Kansas | 3,479 | 3,385 | 2,521 | 72.5 | 74.5 |
| Missouri | 5,945 | 4,907 | 4,061 | 68.3 | 82.8 |
| New Hampshire | 859 | 834 | 620 | 72.2 | 74.3 |
| New Jersey | 11,492 | 11,501 | 7,085 | 61.7 | 61.6 |
| New Mexico | * | * | 136 | — | — |
| New York | 21,618 | 21,377 | 12,370 | 57.2 | 57.9 |
| North Dakota | 624 | 603 | 574 | 92.0 | 95.2 |
| Oklahoma | 5,853 | 5,880 | 4,700 | 80.3 | 79.9 |
| Pennsylvania | * | * | 11,645 | — | — |
| Rhode Island | 1,000 | 943 | 820 | 82.0 | 87.0 |
| Utah | 4,258 | 3,944 | 3,609 | 84.8 | 91.5 |
| West Virginia | 1,707 | 1,636 | 1,286 | 75.3 | 78.6 |
| Wisconsin | 5,884 | 5,477 | 3,781 | 64.3 | 69.0 |

* State did not participate in DLM assessments for these administration years.

In addition to the overall participation rate for DLM assessments declining from 2018–2019 to 2020–2021, we also observed small shifts in the composition of the population across years. Table 4 summarizes the demographic characteristics of the students who participated in the 2017–2018, 2018–2019, and the 2020–2021 administration. There were small percentage point decreases within the distributions of African American students (−4.4, Cohen's *h* = −0.11) who took assessments in 2021, and smaller decreases for students of Hispanic ethnicity (−2.4, Cohen's *h* = −0.06) and English learners (−0.6, Cohen's *h* = −0.03). There was also a notable percentage point increase for white students (5.9, Cohen's *h* = 0.12) taking assessments in 2020–2021, which may be expected given the observed decreases for other racial subgroups.

**Table 4**

*Demographic Characteristics of Participants in 2017–2018, 2018–2019, and 2020–2021*

| | 2017–2018 | | 2018–2019 | | 2020–2021 | | % Point Change | |
|---|---|---|---|---|---|---|---|---|
| Subgroup | *n* | % | *n* | % | *n* | % | 2018 to 2019 | 2019 to 2021 |
| Gender | | | | | | | | |
| Male | 57,604 | 66.6 | 61,279 | 66.6 | 47,412 | 67.7 | -0.1 | 1.1 |
| Female | 28,864 | 33.4 | 30,794 | 33.4 | 22,626 | 32.3 | 0.1 | -1.1 |
| Missing | 3 | 0.0 | 0 | 0.0 | 0 | 0.0 | 0.0 | 0.0 |
| Race | | | | | | | | |
| White | 53,278 | 61.6 | 55,637 | 60.4 | 46,454 | 66.3 | -1.2 | 5.9 |
| African American | 17,319 | 20.0 | 18,814 | 20.4 | 11,200 | 16.0 | 0.4 | -4.4 |
| Two or more races | 8,373 | 9.7 | 10,091 | 11.0 | 7,247 | 10.3 | 1.3 | -0.6 |
| Asian | 4,068 | 4.7 | 4,316 | 4.7 | 2,899 | 4.1 | 0.0 | -0.5 |
| American Indian | 2,655 | 3.1 | 2,564 | 2.8 | 1,738 | 2.5 | -0.3 | -0.3 |
| Native Hawaiian or Pacific Islander | 444 | 0.5 | 468 | 0.5 | 331 | 0.5 | 0.0 | 0.0 |
| Alaska Native | 268 | 0.3 | 183 | 0.2 | 169 | 0.2 | -0.1 | 0.0 |
| Missing | 66 | 0.1 | 0 | 0.0 | 0 | 0.0 | -0.1 | 0.0 |
| Hispanic ethnicity | | | | | | | | |
| Non-Hispanic | 67,858 | 78.5 | 72,596 | 78.8 | 56,857 | 81.2 | 0.4 | 2.3 |
| Hispanic | 18,573 | 21.5 | 19,477 | 21.2 | 13,181 | 18.8 | -0.3 | -2.3 |
| Missing | 40 | 0.0 | 0 | 0.0 | 0 | 0.0 | 0.0 | 0.0 |
| English learning (EL) participation | | | | | | | | |
| Not EL eligible or monitored | 81,449 | 94.2 | 86,848 | 94.3 | 66,487 | 94.9 | 0.1 | 0.6 |
| EL eligible or monitored | 5,022 | 5.8 | 5,225 | 5.7 | 3,551 | 5.1 | -0.1 | -0.6 |

Table 5 displays the counts and proportion of participating students classified to each complexity band for 2017–2018, 2018–2019, and 2020–2021. From 2018–2019 to 2020–2021, a smaller proportion of students were classified to the Foundational Band and Band 3, and a larger proportion were classified to Band 1 and Band 2. However, like in Table 4, these changes were relatively small, with the largest change being only 1.6 percentage points (Cohen's *h* = 0.03).

**Table 5**

*Complexity Band Distribution for 2017–2018, 2018–2019, and 2020–2021*

| | 2017–2018 | | 2018–2019 | | 2020–2021 | | % Point Change | |
|---|---|---|---|---|---|---|---|---|
| Complexity Band | *n* | % | *n* | % | *n* | % | 2018 to 2019 | 2019 to 2021 |
| Foundational | 30,230 | 15.1 | 32,766 | 15.3 | 23,339 | 13.9 | 0.2 | -1.3 |
| Band 1 | 72,637 | 36.3 | 79,912 | 37.2 | 65,137 | 38.8 | 1.0 | 1.6 |
| Band 2 | 73,365 | 36.6 | 78,187 | 36.4 | 61,561 | 36.7 | -0.2 | 0.3 |
| Band 3 | 24,135 | 12.0 | 23,868 | 11.1 | 17,695 | 10.5 | -0.9 | -0.6 |

*Note.* Complexity bands are counted for each subject.

Additional evidence from First Contact survey responses revealed differences in reported attention to different modalities of instruction, with attention decreasing, as shown in Table 6.

**Table 6**

*Percentage Point Difference in Attention to Instruction*

| | Generally sustains attention | | Fleeting attention | | Little or no attention | |
|---|---|---|---|---|---|---|
| Type of instruction | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 |
| Computer-directed | -0.7 | -7.1 | 0.5 | 5.5 | 0.1 | 1.7 |
| Teacher-directed | -0.4 | -4.0 | 0.5 | 3.4 | 0.1 | 0.5 |

During 2020–2021, participating teachers had a median of three students assigned to them in the system and actually administered testlets to a median of two students. For both 2017–2018 and 2018–2019, teachers also had a median of three students assigned to them in the system and administered testlets to a median of three students. Thus, although teachers had similar numbers of students assigned to them as in previous years, they administered testlets to fewer of those students.

We further examined characteristics of students who were enrolled but did not participate. Approximately 20,000 students were enrolled in the DLM system for 2020–2021 but did not respond to any assessment items, compared to about 4,000 students in a typical year. Table 7 summarizes the demographic characteristics of these students. Note that due to small samples sizes in some of the non-participant subgroups, student counts for non-participants have been randomly rounded up or down to the nearest 10 to ensure the data are de-identified (Matthews & Harel, 2011). The non-participant group had higher proportions of students who were African American, of Hispanic ethnicity, English learners, and in the Foundational complexity band compared to the proportions that participated in 2020–2021; however, 4.7% of non-participating students did not have complexity band information available. These demographic groups appear disproportionately less likely to participate in the DLM assessment during the 2020–2021 school year.

**Table 7**

*Demographic Characteristics of Enrolled Participants and Non-Participants in 2020–2021*

| | Participants | | Non-Participants | | |
|---|---|---|---|---|---|
| Subgroup | *n* | % | *n*† | % | % Point Change |
| Gender | | | | | |
| Male | 47,412 | 67.7 | 13,530 | 67.0 | -0.6 |
| Female | 22,626 | 32.3 | 6,640 | 32.9 | 0.6 |
| Non-binary/undesignated | 0 | 0.0 | 10 | <0.1 | 0.0 |
| Race | | | | | |
| White | 46,454 | 66.3 | 10,940 | 54.3 | -12.0 |
| African American | 11,200 | 16.0 | 5,520 | 27.4 | 11.4 |
| Two or more races | 7,247 | 10.3 | 1,200 | 6.0 | -4.4 |
| Asian | 2,899 | 4.1 | 1,500 | 7.4 | 3.3 |
| American Indian | 1,738 | 2.5 | 730 | 3.6 | 1.1 |
| Native Hawaiian or Pacific Islander | 331 | 0.5 | 170 | 0.8 | 0.4 |
| Alaska Native | 169 | 0.2 | 90 | 0.4 | 0.2 |
| Hispanic ethnicity | | | | | |
| Non-Hispanic | 56,857 | 81.2 | 13,840 | 68.6 | -12.6 |
| Hispanic | 13,181 | 18.8 | 6,340 | 31.4 | 12.6 |
| English learning (EL) participation | | | | | |
| Not EL eligible or monitored | 66,487 | 94.9 | 18,100 | 89.7 | -5.2 |
| EL eligible or monitored | 3,551 | 5.1 | 2,070 | 10.3 | 5.2 |
| Complexity Band | | | | | |
| Foundational | 23,339 | 13.9 | 9,950 | 20.8 | 6.9 |
| Band 1 | 65,137 | 38.8 | 16,870 | 35.3 | -3.5 |
| Band 2 | 61,561 | 36.7 | 14,490 | 30.3 | -6.4 |
| Band 3 | 17,695 | 10.5 | 4,240 | 8.9 | -1.7 |
| Missing | 0 | 0.0 | 2,240 | 4.7 | 4.7 |

*Note.* Complexity bands are counted for each subject.

† Values are randomly rounded to the nearest 10.

# 4. Educational Experience of DLM Students

Information about students' educational experience came from teacher survey responses about instruction and assessment and entry of special circumstance codes about students' assessment experience.

The spring 2021 DLM teacher survey collected information about students' educational experience during 2020–2021. Survey questions on instructional experience covered time spent in various instructional settings, scheduling scenarios impacting instruction, the number of 9-week periods the teacher was responsible for the student's instruction, and instructional time as an indicator of opportunity to learn. Survey questions on assessment experience covered testing location, circumstances for students not taking assessments at school, and teachers' perceptions of students' experiences with DLM testlets. Survey responses were optional; the number of survey responses also varied by question due to the spiral assignment of some of the survey blocks.

## 4.1. Instructional Experience

During the COVID-19 pandemic, students may have been instructed in a variety of instructional settings. Teachers were asked to report the percentage of time students spent in each instructional setting, as shown in Table 8. Teachers indicated that 47% of students ($n$ = 17,272) spent more than 75% of their instructional time in school and that 21% of students ($n$ = 7,587) spent 50%–75% of their instructional time in school. More than half of responses indicated at least some time spent in-home with direct instruction from the teacher (one-on-one and/or as a class). Teachers reported that 31% of students received some amount of instruction in their home with their family member providing the instruction and 6% of students received no in-school instruction during the year.

To evaluate whether student groups were potentially differentially impacted, we examined instructional setting for subgroup differences for gender, race, ethnicity, first language, and complexity band. We specifically compared the percentage of students spending 50% or less of their time and greater than 50% of their time in each instructional setting across subgroups. Table 9 shows that white students were more likely to spend more than half of their time in school compared to the other racial subgroups. Similarly, only 62% ($n$ = 3,733) of Hispanic students and 62% ($n$ = 1,190) of EL eligible or monitored students spent more than 50% of the year in school compared to 70% of non-Hispanic students ($n$ = 21,126, Cohen's $h$ = −0.17) and 69% of non-EL eligible students ($n$ = 23,669, Cohen's $h$ = −0.16). African American, Asian, Hispanic, and EL eligible students were more likely than white students and non-EL eligible students to spend more than half of their instructional time in the home, either with direct remote one-on-one instruction with the teacher or remote group or whole class instruction. Students classified to the Foundational complexity band (66%, $n$ = 9,826) were only slightly less likely to spend more than 50% of time at school compared to students at Band 1 (70%, $n$ = 28,556, Cohen's $h$ = −0.09), Band 2 (70%, $n$ = 26,981, Cohen's $h$ = −0.09), and Band 3 (69%, $n$ = 8,131, Cohen's $h$ = −0.07). There were no gender differences in the time spent in each instructional setting.

**Table 8**

*Percentage of Time Spent in Each Instructional Setting*

| Instructional setting | None | | 1%–25% | | 26%–50% | | 51%–75% | | 76%–100% | | Unknown | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| In school | 2,263 | 6.2 | 3,856 | 10.5 | 5,025 | 13.7 | 7,587 | 20.7 | 17,272 | 47.2 | 591 | 1.6 |
| In-home, direct instruction with teacher (remotely, 1:1) | 13,974 | 40.8 | 12,061 | 35.2 | 3,644 | 10.6 | 1,892 | 5.5 | 1,443 | 4.2 | 1,245 | 3.6 |
| In-home, direct instruction with teacher (remotely, group or whole class) | 12,300 | 35.3 | 12,107 | 34.7 | 4,460 | 12.8 | 2,549 | 7.3 | 2,308 | 6.6 | 1,161 | 3.3 |
| In-home, teacher present in home | 30,011 | 89.7 | 923 | 2.8 | 446 | 1.3 | 327 | 1.0 | 366 | 1.1 | 1,378 | 4.1 |
| In-home, family member providing instruction | 20,401 | 60.2 | 7,102 | 21.0 | 1,601 | 4.7 | 880 | 2.6 | 897 | 2.6 | 2,987 | 8.8 |
| Absent | 21,548 | 65.0 | 7,473 | 22.5 | 893 | 2.7 | 497 | 1.5 | 369 | 1.1 | 2,365 | 7.1 |
| Other | 22,886 | 80.5 | 618 | 2.2 | 251 | 0.9 | 205 | 0.7 | 282 | 1.0 | 4,177 | 14.7 |

**Table 9**

*Time Spent in Each Instructional Setting by Subgroup*

| | In school | | In-home, direct instruction with teacher (remotely, 1:1) | | In-home, direct instruction with teacher (remotely, group or whole class) | | In-home, teacher present in home | | In-home, family member providing instruction | | Absent | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Subgroup | <50% | >50% | <50% | >50% | <50% | >50% | <50% | >50% | <50% | >50% | <50% | >50% |
| Gender | | | | | | | | | | | | |
| Female | 30.6 | 69.4 | 90.4 | 9.6 | 85.3 | 14.7 | 97.6 | 2.4 | 94.2 | 5.8 | 97.1 | 2.9 |
| Male | 31.1 | 68.9 | 89.7 | 10.3 | 85.7 | 14.3 | 98.0 | 2.0 | 94.2 | 5.8 | 97.2 | 2.8 |
| Race | | | | | | | | | | | | |
| African American | 48.3 | 51.7 | 84.4 | 15.6 | 73.6 | 26.4 | 96.9 | 3.1 | 92.7 | 7.3 | 94.9 | 5.1 |
| Alaska Native | 35.0 | 65.0 | 93.6 | 6.4 | 90.1 | 9.9 | 100.0 | 0.0 | 94.4 | 5.6 | 92.5 | 7.5 |
| American Indian | 28.3 | 71.7 | 88.7 | 11.3 | 88.6 | 11.4 | 98.2 | 1.8 | 90.3 | 9.7 | 96.5 | 3.5 |
| Asian | 36.5 | 63.5 | 85.6 | 14.4 | 81.4 | 18.6 | 97.4 | 2.6 | 93.4 | 6.6 | 96.5 | 3.5 |
| Native Hawaiian or Pacific Islander | 34.4 | 65.6 | 88.6 | 11.4 | 88.7 | 11.3 | 97.7 | 2.3 | 95.3 | 4.7 | 95.0 | 5.0 |
| Two or more races | 40.5 | 59.5 | 88.4 | 11.6 | 79.0 | 21.0 | 97.9 | 2.1 | 93.8 | 6.2 | 96.4 | 3.6 |
| White | 25.3 | 74.7 | 91.7 | 8.3 | 89.5 | 10.5 | 98.0 | 2.0 | 94.9 | 5.1 | 97.9 | 2.1 |
| Hispanic ethnicity | | | | | | | | | | | | |
| Hispanic | 37.7 | 62.3 | 88.1 | 11.9 | 79.7 | 20.3 | 97.4 | 2.6 | 94.5 | 5.5 | 96.0 | 4.0 |
| Non-Hispanic | 29.6 | 70.4 | 90.3 | 9.7 | 86.8 | 13.2 | 97.9 | 2.1 | 94.2 | 5.8 | 97.4 | 2.6 |
| English learning (EL) participation | | | | | | | | | | | | |
| EL eligible or monitored | 38.1 | 61.9 | 87.1 | 12.9 | 78.9 | 21.1 | 96.9 | 3.1 | 94.4 | 5.6 | 96.0 | 4.0 |
| Not EL eligible or monitored | 30.6 | 69.4 | 90.1 | 9.9 | 86.0 | 14.0 | 97.9 | 2.1 | 94.2 | 5.8 | 97.3 | 2.7 |
| Complexity band | | | | | | | | | | | | |
| Foundational | 34.5 | 65.5 | 90.1 | 9.9 | 88.5 | 11.5 | 96.7 | 3.3 | 92.0 | 8.0 | 96.0 | 4.0 |
| Band 1 | 30.3 | 69.7 | 90.0 | 10.0 | 87.3 | 12.7 | 97.8 | 2.2 | 93.7 | 6.3 | 96.8 | 3.2 |
| Band 2 | 30.4 | 69.6 | 89.8 | 10.2 | 84.1 | 15.9 | 98.2 | 1.8 | 95.1 | 4.9 | 97.8 | 2.2 |
| Band 3 | 31.1 | 68.9 | 89.3 | 10.7 | 80.3 | 19.7 | 97.9 | 2.1 | 95.8 | 4.2 | 97.8 | 2.2 |

Teachers reported a number of instructional scheduling scenarios that impacted students taking the DLM assessments. Teachers were asked which scheduling scenarios applied to the student in 2020–2021, as shown in Table 10. Most responses (71.6%, *n* = 27,019) indicated changes between remote and in-person learning at least once during the school year. Teachers reported that around a third of students had a delayed start to the school year and/or extended school year through summer. It is unknown how many of those students typically participate in an extended school year.

**Table 10**

*Student Scheduling Scenarios*

| | Yes | | No | | Unknown | |
|---|---|---|---|---|---|---|
| Scheduling scenario | *n* | % | *n* | % | *n* | % |
| Delayed start of the 2020–2021 school year | 10,536 | 28.6 | 25,337 | 68.7 | 1,020 | 2.8 |
| Lengthened spring 2021 semester | 1,673 | 4.6 | 33,415 | 92.0 | 1,236 | 3.4 |
| Extended school year through summer 2021 | 13,433 | 36.7 | 21,010 | 57.3 | 2,193 | 6.0 |
| Change(s) between remote and in-person learning during the 2020–2021 school year | 27,019 | 71.6 | 10,022 | 26.6 | 696 | 1.8 |

Due to changes in instructional settings and schedules, students may have had more than one teacher for the full school year. Teachers were asked about the number of 9-week periods of the school year for which they were responsible for their student's instruction, as shown in Table 11. While teachers mostly reported that they were responsible for the student's instruction for the full year (i.e., four or more 9-week periods; 72.4%, *n* = 27,667), for 20.5% of students, their teacher reported they were not responsible for instruction for the full year. This could be the result of students transferring schools during the year. Alternatively, these results could indicate the student had more than one teacher (e.g., when transitioning from remote to in-person learning), or could correspond to survey responses that a parent provided instruction for part of the year. Additional research is needed to investigate this further.

**Table 11**

*Number of 9-week Periods Teacher Responsible for Student's Instruction*

| Number of 9-week periods | *n* | % |
|---|---|---|
| One | 1,349 | 3.5 |
| Two | 1,835 | 4.8 |
| Three | 4,658 | 12.2 |
| Four or more | 27,667 | 72.4 |
| Not applicable* | 2,682 | 7.0 |

* This option may have been selected if the individual completing the survey was not the student's teacher and not responsible for the student's instruction.

Disruptions caused by the COVID-19 pandemic may have also impacted students' total amount of instructional time and their opportunity to learn academic content. Table 12, Table 13, Table 14, and Table 15 show trends from 2017–2018 to 2020–2021 in the total number of hours of instruction for ELA, mathematics, science core ideas, and science and engineering practices, respectively. In all topics, a larger percentage of students received 0–5 hours of

instruction and a smaller percentage received more than 20 hours of instruction in 2020–2021 compared to 2017–2018 and 2018–2019, including for untested topics. When interpreting these changes, please note that in 2020–2021, teachers completed these items for a larger number of students compared to 2017–2018 and 2018–2019 because in 2020–2021 the teacher survey included a smaller number of spiraled forms. In addition, sample size across years varied by subject due to the survey containing spiraled forms; the ELA and mathematics questions received 5,000–7,500 responses per item per year while the science questions received 1,900–2,800 responses. In science, the response options changed from 2017–2018 to 2018–2019.[2]

To further explore students' opportunity to learn during the pandemic, we developed a set of indicators on breadth and depth of instruction in ELA, mathematics, and science using the data shown in Tables 12–15. Breadth of instruction in each subject was computed as the number of topics (i.e., conceptual areas, core ideas, or science and engineering practices) in which teachers indicated 6 or more hours of instruction during the school year. Depth of instruction was computed as the number of topics for which teachers indicated more than 15 hours of instruction during the school year.[3]

Table 16 shows average breadth and depth in each subject area in 2017–2018, 2018–2019, and 2020–2021. Due to the acknowledged impacts of the pandemic on learning, we expected to see a decline in breadth and depth of instruction; in all subjects, there were small declines with effect sizes (Cohen's *d*) ranging from 0.12 to 0.23.

[2] In 2017–2018, the response options were *none*, *1–10 hours*, *11–20 hours*, *21–30 hours*, and *more than 30 hours*. In 2018–2019 and 2020–2021, the response options were *0–5 hours*, *5–10 hours*, *11–15 hours*, *16–20 hours*, and *more than 20 hours*.

[3] Because the first response option is *0–5 hours* and teachers selecting this option may have provided no instruction, this option was not included in the breadth indicator. Additionally, because the response options differed, breadth for science in 2017–2018 was defined as the number of topics with at least 1 hour of instruction and depth as the number of topics with at least 11 hours of instruction.

**Table 12**

*Hours of Instructional Time Spent on English Language Arts Conceptual Areas*

| | 0–5 (%) | | | 6–10 (%) | | | 11–15 (%) | | | 16–20 (%) | | | >20 (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Conceptual area | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 |
| Determine critical elements of text | 23.1 | 23.7 | 31.4 | 16.2 | 14.9 | 17.0 | 12.6 | 14.1 | 13.5 | 14.9 | 14.8 | 13.0 | 33.3 | 32.6 | 25.2 |
| Construct understandings of text | 16.3 | 17.3 | 21.7 | 15.9 | 14.4 | 16.1 | 13.4 | 13.8 | 14.4 | 15.6 | 16.3 | 15.4 | 38.7 | 38.3 | 32.5 |
| Integrate ideas and information from text | 18.9 | 19.5 | 25.4 | 17.0 | 15.6 | 18.1 | 14.1 | 15.0 | 15.9 | 17.0 | 17.0 | 15.5 | 32.9 | 32.9 | 25.1 |
| Use writing to communicate | 23.7 | 24.2 | 29.3 | 15.9 | 15.5 | 17.4 | 13.4 | 14.2 | 14.2 | 15.1 | 14.5 | 14.7 | 32.0 | 31.6 | 24.3 |
| Integrate ideas and information in writing | 26.4 | 27.1 | 34.6 | 16.2 | 15.9 | 17.9 | 14.4 | 14.6 | 14.7 | 15.2 | 15.3 | 13.8 | 27.8 | 27.0 | 19.0 |
| Use language to communicate with others | 9.8 | 10.4 | 12.6 | 11.3 | 10.1 | 10.5 | 11.2 | 12.1 | 12.9 | 15.0 | 15.6 | 16.2 | 52.8 | 51.7 | 47.8 |
| Clarify and contribute in discussion | 17.5 | 18.1 | 22.5 | 15.4 | 14.6 | 15.7 | 14.5 | 14.5 | 15.3 | 16.8 | 16.8 | 17.1 | 35.8 | 36.0 | 29.5 |
| Use sources and information | 28.1 | 29.2 | 39.2 | 18.3 | 17.1 | 19.5 | 15.2 | 15.4 | 14.5 | 15.2 | 14.9 | 11.7 | 23.2 | 23.4 | 15.1 |
| Collaborate and present ideas | 26.5 | 27.9 | 37.6 | 18.8 | 17.3 | 19.6 | 15.9 | 15.4 | 14.9 | 15.7 | 15.7 | 12.2 | 23.0 | 23.7 | 15.8 |

*Note.* Only the first five conceptual areas listed in this table are measured by the DLM assessment.

**Table 13**

*Hours of Instructional Time Spent on Mathematics Conceptual Areas*

| | 0–5 (%) | | | 6–10 (%) | | | 11–15 (%) | | | 16–20 (%) | | | >20 (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Conceptual area | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 |
| Understand number structures (counting, place value, fraction) | 15.0 | 15.4 | 17.9 | 13.9 | 13.2 | 14.5 | 11.7 | 12.1 | 12.8 | 14.8 | 15.0 | 16.2 | 44.7 | 44.2 | 38.6 |
| Compare, compose, and decompose numbers and steps | 26.5 | 26.3 | 32.5 | 16.8 | 16.8 | 17.9 | 14.3 | 14.5 | 14.6 | 15.1 | 15.3 | 15.2 | 27.3 | 27.1 | 19.8 |
| Calculate accurately and efficiently using simple arithmetic operations | 22.0 | 23.0 | 25.6 | 13.0 | 12.1 | 12.8 | 11.6 | 12.4 | 12.4 | 14.5 | 15.0 | 15.9 | 39.0 | 37.4 | 33.3 |
| Understand and use geometric properties of two- and three-dimensional shapes | 32.5 | 33.3 | 40.0 | 21.1 | 21.6 | 22.3 | 16.7 | 16.0 | 16.2 | 15.0 | 14.4 | 12.2 | 14.7 | 14.6 | 9.3 |
| Solve problems involving area, perimeter, and volume | 52.1 | 52.8 | 61.1 | 15.6 | 16.1 | 16.4 | 12.2 | 11.7 | 10.1 | 10.4 | 10.1 | 7.3 | 9.8 | 9.3 | 5.2 |
| Understand and use measurement principles and units of measure | 34.5 | 35.1 | 42.2 | 23.1 | 22.8 | 23.7 | 16.6 | 16.4 | 15.6 | 12.8 | 13.2 | 10.4 | 13.0 | 12.4 | 8.1 |
| Represent and interpret data displays | 34.2 | 34.6 | 41.8 | 20.0 | 20.5 | 21.2 | 18.0 | 17.0 | 15.8 | 13.8 | 13.5 | 11.7 | 14.0 | 14.3 | 9.4 |
| Use operations and models to solve problems | 28.0 | 29.1 | 33.6 | 16.7 | 15.9 | 17.4 | 15.3 | 15.3 | 15.1 | 15.9 | 16.9 | 15.6 | 24.0 | 22.8 | 18.3 |
| Understand patterns and functional thinking | 21.9 | 22.8 | 28.8 | 20.5 | 20.0 | 21.5 | 19.3 | 18.8 | 18.7 | 16.6 | 17.0 | 16.0 | 21.7 | 21.3 | 15.1 |

**Table 14**

*Hours of Instructional Time Spent on Science Core Ideas*

| | 0–10 (%) | | | 11–20 (%) | | | >20 (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| Core idea | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 |
| Physical Science | | | | | | | | | |
| Matter and Its Interactions | 64.3 | 66.2 | 77.5 | 19.5 | 23.2 | 17.5 | 16.2 | 10.6 | 5.1 |
| Motion and Stability: Forces and Interactions | 66.4 | 70.6 | 78.7 | 19.3 | 20.8 | 16.9 | 14.3 | 8.6 | 4.4 |
| Energy | 65.7 | 67.8 | 77.1 | 19.0 | 22.4 | 18.3 | 15.3 | 9.8 | 4.6 |
| Life Science | | | | | | | | | |
| From Molecules to Organisms: Structure and Processes | 67.3 | 72.6 | 76.6 | 17.6 | 19.4 | 18.5 | 15.1 | 8.0 | 4.9 |
| Ecosystems: Interactions, Energy, and Dynamics | 59.6 | 62.3 | 67.5 | 20.5 | 26.1 | 25.0 | 19.9 | 11.7 | 7.5 |
| Heredity: Inheritance and Variation of Traits | 73.1 | 76.7 | 79.3 | 14.3 | 16.5 | 16.8 | 12.6 | 6.8 | 3.9 |
| Biological Evolution: Unity and Diversity | 71.3 | 75.2 | 78.0 | 15.8 | 17.6 | 17.7 | 12.8 | 7.2 | 4.2 |
| Earth and Space Science | | | | | | | | | |
| Earth's Place in the Universe | 64.0 | 67.5 | 70.7 | 19.3 | 23.3 | 23.4 | 16.7 | 9.1 | 5.9 |
| Earth's Systems | 63.4 | 67.5 | 71.1 | 19.8 | 23.4 | 22.5 | 16.8 | 9.2 | 6.4 |
| Earth and Human Activity | 61.1 | 64.7 | 67.1 | 20.2 | 25.3 | 25.6 | 18.7 | 10.0 | 7.3 |

**Table 15**

*Hours of Instructional Time Spent on Science and Engineering Practices*

| | 0–10 (%) | | | 11–20 (%) | | | >20 (%) | | |
|---|---|---|---|---|---|---|---|---|---|
| Practice | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 | 2018 | 2019 | 2021 |
| Developing and Using Models | 71.1 | 71.5 | 77.0 | 15.5 | 20.6 | 17.6 | 13.5 | 7.9 | 5.4 |
| Planning and Carrying Out Investigations | 64.0 | 67.9 | 71.8 | 19.6 | 23.3 | 21.3 | 16.4 | 8.8 | 6.9 |
| Analyzing and Interpreting Data | 60.4 | 62.5 | 68.3 | 21.2 | 25.8 | 23.7 | 18.4 | 11.7 | 8.0 |
| Using Mathematics and Computational Thinking | 57.3 | 58.8 | 67.6 | 19.0 | 25.2 | 22.6 | 23.7 | 16.0 | 9.8 |
| Constructing Explanations and Designing Solutions | 68.6 | 69.7 | 75.4 | 17.6 | 21.7 | 19.5 | 13.8 | 8.6 | 5.0 |
| Engaging in Argument From Evidence | 73.1 | 74.0 | 79.4 | 15.1 | 18.8 | 16.3 | 11.8 | 7.2 | 4.3 |
| Obtaining, Evaluating, and Communicating Information | 60.4 | 62.2 | 67.0 | 19.2 | 25.6 | 22.9 | 20.5 | 12.2 | 10.0 |

**Table 16**

*Breadth and Depth of Instruction in 2017–2018, 2018–2019, and 2020–2021*

| | | 2017–2018 | | | 2018–2019 | | | 2020–2021 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Subject | Max | *n* | Mean | SD | *n* | Mean | SD | *n* | Mean | SD |
| Breadth | | | | | | | | | | |
| ELA | 9 | 5,448 | 7.0 | 3.0 | 5,669 | 7.0 | 3.1 | 7,523 | 6.4 | 3.2 |
| Mathematics | 9 | 5,429 | 6.3 | 3.1 | 5,723 | 6.2 | 3.2 | 7,975 | 5.7 | 3.2 |
| Science Core Ideas | 10 | 1,908 | 7.3 | 3.3 | 2,259 | 5.1 | 4.1 | 2,884 | 4.5 | 4.0 |
| Science and Engineering Practices | 7 | 1,900 | 5.3 | 2.4 | 2,254 | 3.9 | 3.0 | 2,871 | 3.4 | 2.9 |
| Depth | | | | | | | | | | |
| ELA | 9 | 5,448 | 4.4 | 3.7 | 5,669 | 4.3 | 3.7 | 7,523 | 3.6 | 3.4 |
| Mathematics | 9 | 5,429 | 3.3 | 3.3 | 5,723 | 3.3 | 3.2 | 7,975 | 2.7 | 2.9 |
| Science Core Ideas | 10 | 1,908 | 3.4 | 3.9 | 2,259 | 1.7 | 3.2 | 2,884 | 1.3 | 2.6 |
| Science and Engineering Practices | 7 | 1,900 | 2.4 | 2.8 | 2,254 | 1.4 | 2.4 | 2,871 | 1.0 | 2.0 |

Teachers also provided overall ratings on the number of hours per week students engaged in academic instruction. Table 17 shows declines from 2017–2018 to 2020–2021 in the percentage of students receiving between 6 and 20 hours of instruction (Cohen's *h* ranging from −0.10 to −0.02) and an increase in the percentage receiving more than 20 hours of weekly instruction (Cohen's *h* ranging from 0.08 to 0.15). This finding was unexpected and should be explored further. Note that the response options for this survey question changed on the 2020–2021 survey, which may have contributed to the unexpected findings.

**Table 17**

*Trends in Time Spent Engaging Student in Academic Instruction*

| | 2017–2018 | | 2018–2019 | | 2020–2021 | |
|---|---|---|---|---|---|---|
| Hours per week | *n* | % | *n* | % | *n* | % |
| None | 398 | 0.7 | 383 | 0.7 | 370 | 1.0 |
| 1 to 5 | 5,840 | 10.4 | 4,913 | 9.3 | 3,682 | 9.7 |
| 6 to 10 | 7,963 | 14.2 | 6,968 | 13.1 | 4,406 | 11.6 |
| 11 to 15 | 8,188 | 14.6 | 7,418 | 14.0 | 4,477 | 11.8 |
| 15 to 20 | 13,495 | 24.1 | 12,308 | 23.2 | 7,307 | 19.2 |
| More than 20[†] | 20,090 | 35.9 | 21,101 | 39.7 | 17,830 | 46.9 |

[†] On the 2020–2021 survey the response options were 21–30 hours (*n* = 11,559; 30.4%) and more than 30 hours (*n* = 6,271; 16.5%).

To explore whether these changes were observed within students or were potentially a product of sampling, we examined response records for students who participated in DLM assessments and had teacher survey responses in both 2018–2019 and 2020–2021. Table 18 shows the total amount of academic instruction per week for students whose teachers responded to this survey item in both 2018–2019 and 2020–2021 (*n* = 12,443). Responses may have been submitted by the same or different teacher in 2018–2019 and 2020–2021. Approximately 37% of students received more academic instruction in 2020–2021 compared to 2018–2019, 31% received less instruction, and 32% received the same amount of instruction. However, this survey question does not indicate how instructional time was spent. For example, teachers may have spent more time on a narrower range of academic content. This would result in fewer conceptual areas receiving instruction, which is consistent with the changes in breadth and depth of instruction shown in Table 16. Future research is planned to examine these findings further.

**Table 18**

*Hours Spent Engaging Student in Academic Instruction for Students with Survey Data in 2018–2019 and 2020–2021*

| | 2020–2021 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 5 or fewer | | 6 to 10 | | 11 to 15 | | 16 to 20 | | More than 20 | |
| 2018–2019 | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| 5 or fewer | 230 | 1.8 | 179 | 1.4 | 136 | 1.1 | 183 | 1.5 | 393 | 3.2 |
| 6 to 10 | 224 | 1.8 | 251 | 2.0 | 235 | 1.9 | 304 | 2.4 | 636 | 5.1 |
| 11 to 15 | 177 | 1.4 | 227 | 1.8 | 238 | 1.9 | 357 | 2.9 | 728 | 5.9 |
| 16 to 20 | 257 | 2.1 | 297 | 2.4 | 383 | 3.1 | 635 | 5.1 | 1,419 | 11.4 |
| More than 20 | 421 | 3.4 | 473 | 3.8 | 485 | 3.9 | 932 | 7.5 | 2,643 | 21.2 |

On the 2020–2021 survey, teachers were asked to describe the most important factors that impacted students' instruction during the pandemic. Many of the responses centered around the student's learning location (remote or in-person). While some teachers stated that their students did well in both remote and in-person learning, many teachers stated that students struggled with remote learning. One teacher wrote, "Remote learning was not effective for this student. Without control of the environment and reinforcement schedule it was extremely difficult to engage this student in remote learning."

Factors that impacted instruction during remote learning included student proclivity and willingness to work and stay focused, access to materials, access to and use of technology, access to reliable internet, home environment, and amount of parental support. One teacher noted, "Remote learning was very difficult for the student to attend to without constant adult supervision in the home, which was difficult for parents to provide, given work schedules and other children to assist." Furthermore, some teachers noted that remote instruction proved especially difficult for their students who do not communicate using speech or their students who are blind or have visual impairments. One teacher noted, "This student is non-verbal so the 7 weeks of remote learning was difficult for her. She has a communication device but needs help using it." Another teacher noted, "Remote learning was frustrating due to the fact that the student is legally blind and unable to access the [online] learning without adult help."

Many teachers noted that the ability to attend school in-person was the most important factor impacting student instruction. One teacher stated, "Being in person 99% of the school year has positively impacted the student's response to academic instruction." Another teacher wrote, "It was beneficial for this student to have mostly in person instruction. She has a difficult time working virtually." Additionally, many teachers indicated that the constant switch between in-person and remote instruction significantly impacted learning this year. One teacher stated, "All of the changes from in-person learning to hybrid or fully remote have taken their toll on this student." Teachers also responded that consistency in relation to schedules, structures, and routines most significantly impacted instruction. Many teachers also indicated that absences and lack of attendance, school closures, quarantines, and lack of social interaction were other important factors that impacted instruction.

Teachers also identified COVID-19 safety protocols, mask-wearing, and social distancing mandates as important factors that impacted student instruction. One teacher stated, "It was difficult for the student to hear/see formed words while wearing the mask… Peer modeling/instruction/partner work was also impacted due to having to keep 6 feet apart." Another teacher stated, "He didn't understand the concept of wearing his mask and why he couldn't hug his peers or sit near them." Some teacher responses also centered around certain factors that could have been affected by the pandemic, such as student focus, attention, engagement, motivation, direction-following, behavior, anxiety, stress, and mental health. One teacher wrote, "This student had trouble with her motivation to work, focus, and try her best on the tests." Another teacher stated, "Student's anxiety hindered him this year." A few teachers also noted that their students faced economic hardship, including parental joblessness and homelessness. One teacher stated, "This student was impacted by the COVID-19 virus. His family lost their jobs, home, and car… This had an impact on his education." Some teachers did note other important factors not necessarily related to COVID-19, such as following Individualized Education Plans, language barriers, and staffing issues. Some teachers also noted that a student's physical health (including COVID-19 infections) impacted their instruction. Additionally, teachers also responded that a student's disability and a students' communication skills, comprehension skills, and cognitive skills were important factors that impacted instruction.

## 4.2. Assessment Experience

Students taking DLM assessments respond to a series of testlets during an assessment window, with teachers deciding when to administer each testlet. To further understand students' educational context during the 2020–2021 academic year, we examined the use of special circumstance codes that some states use to indicate the reason a student did not test. We also examined the timing of testlet administration during the assessment windows compared

to prior years, as well as the extent to which students completed all assessments, recognizing that disruptions to instruction and school schedules may have impacted students' opportunity to respond to all assessments. Timing of testlet administration and blueprint coverage findings are summarized for each assessment model. Finally, we summarize teacher survey responses pertaining to assessment administration.

### *4.2.1. Special Circumstance Codes*

The frequency in the use of special circumstance codes provided some information about student experience during the pandemic. During the 2020–2021 operational year, 12 states gave the option of entering one or more of the following special circumstance codes related to COVID-19: 1. Student could not test due to COVID-19; 2. Teacher administered the assessment remotely; and 3. Non-teacher administered. The first code ("Student could not test due to COVID-19") does not necessarily indicate that the student contracted COVID-19, but rather could indicate a variety of possibilities (e.g., schools closed due to an outbreak and testing was unavailable). Out of 12 states making the codes available, 8,144 students from 10 states had at least one special circumstance code related to COVID-19 and remote assessment. Of those students, 8,127 students (99%) could not test due to COVID-19, 16 students (<1%) had their teacher administer the assessment remotely, and five students (<1%) were administered the assessment by someone other than their teacher. Of the 8,127 students with a code indicating they could not test due to COVID-19, 91% did not complete any testlets across subjects; the remaining 9% completed at least one testlet in at least one subject. Because these codes were only available in a subset of states, and were only entered for a small subset of students, they provide some limited information about a subset of students' educational experience.

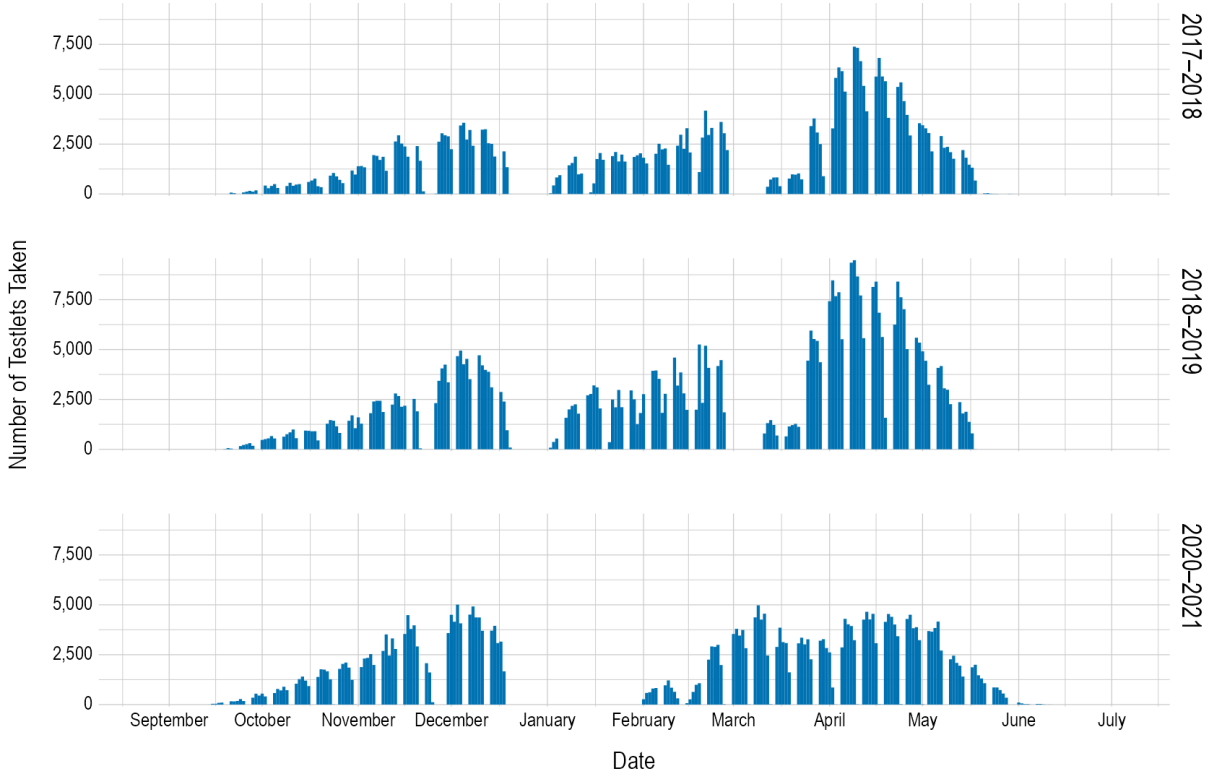
### *4.2.2. Instructionally Embedded Model Administration Data*

#### 4.2.2.1. Timing of Testlet Administration

Figure 1 shows the distribution of the number of testlets taken per day by students in the Instructionally Embedded model. The distribution of the number of testlets completed per day shows the maximum volume of testlets completed per day was reduced in 2020–2021 compared to 2017–2018 and 2018–2019. In 2020–2021, the most testlets taken in a day was 5,010, compared to 7,385 in 2017–2018 and 9,479 in 2018–2019. The fall patterns were similar across years, but the spring pattern differed in 2020–2021. The gap in testing in January 2021 occurred because the spring Instructionally Embedded administration window opened in February and testlets were not available.

**Figure 1**

*Distribution of Testlets Taken per Day for Students in the Instructionally Embedded Model*




### 4.2.2.2. Blueprint Coverage

In years prior to 2019–2020, students in the Instructionally Embedded model were expected to test on the full blueprint in the fall and on a subset of EEs in the spring window. Beginning in 2019–2020, students were expected to test on the full blueprint in both the fall and spring windows. To meet blueprint coverage, students were expected to complete between 6 and 11 testlets, depending on grade and subject.

We examined the percentage of students in three blueprint coverage categories (met, not met, exceeded) for the 2017–2018, 2018–2019, and 2020–2021 administrations, excluding students who did not complete any testlets. In Table 19, we combined student's blueprint coverage for the full administration year. That is, a student with a "Met" designation may not have met coverage requirements in any single window, but did meet all requirements when combining data across the fall and spring windows. Across years, there was a significant decrease in the percentage of students not meeting blueprint requirements and a significant increase in the percentage of students meeting and/or exceeding blueprint requirements. Because the majority of students met blueprint coverage in both the fall and spring testing windows individually, the percentage of students exceeding blueprint requirements for the entire year showed significant increases. In 2020–2021, only 5% or fewer students did not meet all blueprint coverage expectations.

Recall that as part of the transition to two fully embedded windows with teacher selection, the new Instruction and Assessment Planner was introduced to help teachers identify which blueprint requirements had and had not been met. Additionally, teachers were able to choose EEs in both the fall and the spring window. By expecting blueprint coverage in each of the two windows, students had more opportunities to meet or exceed blueprint coverage than in

prior years.

**Table 19**

*Instructionally Embedded Blueprint Coverage in 2017–2018, 2018–2019, and 2020–2021*

| | 2017–2018 | | 2018–2019 | | 2020–2021 | |
|---|---|---|---|---|---|---|
| Category | *n* | % | *n* | % | *n* | % |
| English Language Arts | | | | | | |
| Not met | 4,856 | 41.1 | 3,658 | 24.4 | 485 | 4.0 |
| Met | 4,081 | 34.5 | 6,800 | 45.4 | 4,868 | 40.6 |
| Exceeded | 2,876 | 24.3 | 4,520 | 30.2 | 6,633 | 55.3 |
| Met or exceeded | 6,957 | 58.9 | 11,320 | 75.6 | 11,501 | 96.0 |
| Mathematics | | | | | | |
| Not met | 4,906 | 42.7 | 3,112 | 21.2 | 633 | 5.4 |
| Met | 3,790 | 33.0 | 7,010 | 47.9 | 4,703 | 40.1 |
| Exceeded | 2,797 | 24.3 | 4,526 | 30.9 | 6,378 | 54.4 |
| Met or exceeded | 6,587 | 57.3 | 11,536 | 78.8 | 11,081 | 94.6 |

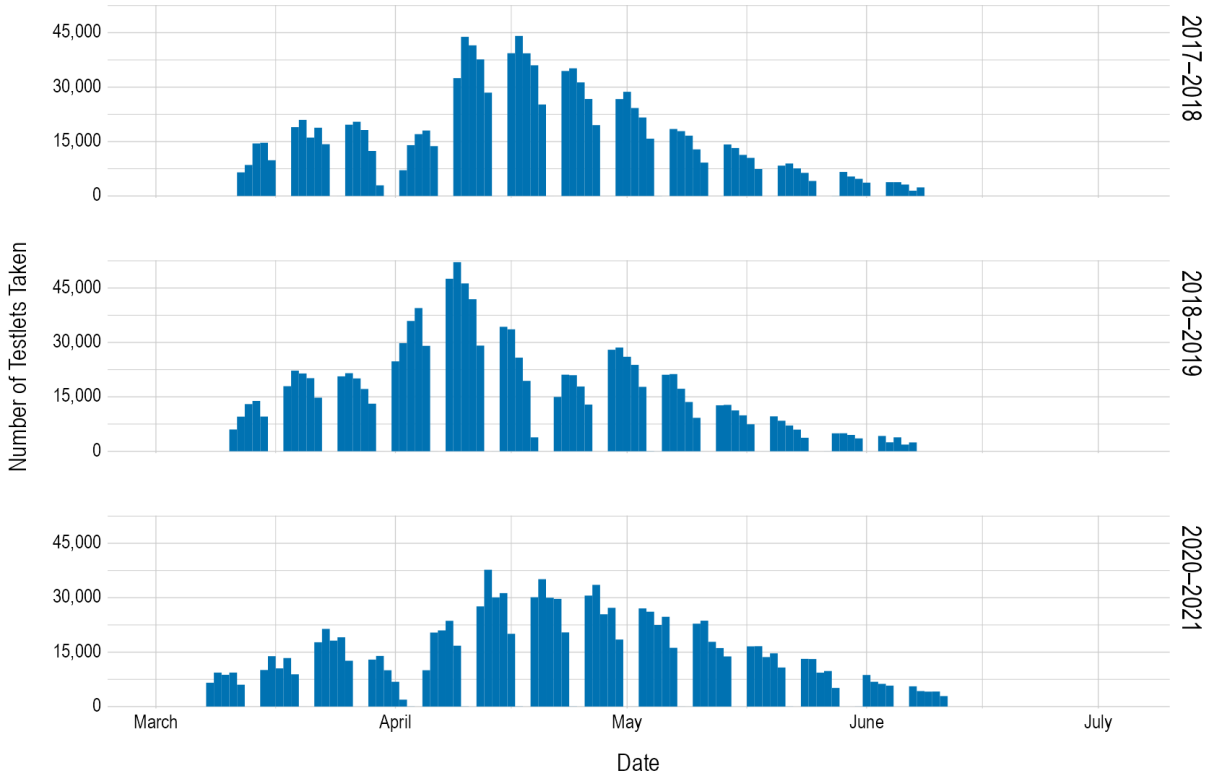
## *4.2.3. Year-End Model Administration Data*

### 4.2.3.1. Timing of Testlet Administration

Figure 2 shows the distribution of the number of testlets taken per day by students in the Year-End model. The distribution of the number of testlets completed per day for the Year-End model shows the maximum volume of testlets completed per day was reduced in 2020–2021 compared to 2017–2018 and 2018–2019. In 2020–2021, the most testlets taken in a day was 37,710, compared to 44,113 in 2017–2018 and 52,118 in 2018–2019.

**Figure 2**

*Distribution of Testlets Taken per Day for Students in the Year-End Model*




#### 4.2.3.2. Blueprint Coverage

In 2020–2021, students in the Year-End model were expected to complete between 6 and 14 testlets, depending on grade and subject, to demonstrate full blueprint coverage. This compared to between 8 and 18 testlets in 2018–2019. The vast majority of students tested on all of the required EEs during the 2020–2021 administration. This is consistent with the 2017–2018 and 2018–2019 administrations. Of students who tested, more than 97% in ELA and mathematics and more than 95% in science tested on all required EEs in each of the 2017–2018, 2018–2019, and 2020–2021 administrations.

### *4.2.4. Survey Data on Assessment Administration*

Disruptions to instruction may have also impacted students' experience taking assessments. Teachers were asked where their student took DLM assessments this year; because students complete multiple testlets, teachers were able to select multiple responses for this question. However, teachers selected more than one response for only 0.4% (*n* = 147) of students. For the vast majority of students, teachers responded that DLM assessments were administered to the student at school (95.4%, *n* = 36,488), as shown in Table 20.

**Table 20**

*DLM Assessment Administration Setting*

| Setting | *n* | % |
|---|---|---|
| At school | 36,488 | 95.4 |
| At home | 658 | 1.7 |
| Testing facility not at school | 227 | 0.6 |
| Other | 190 | 0.5 |
| Not applicable[†] | 857 | 2.2 |

[†] Teachers could select more than one response. The not applicable response was available for teachers to select if the student did not take any DLM assessments. However, the number of responses to this option does not align with other data on the number of students not testing for whom the teacher completed a survey.

The survey also included items that collected information about test administration for students who took any DLM assessments off site, which was allowed under the DLM remote administration guidance so long as a trained test administrator administered the assessment in person. Most students did not take any DLM assessments remotely; after filtering the data to those who did, very few teachers indicated disruptions to their students' standard administration experience. As shown in Table 21, very few teachers indicated disruptions to their students' standard administration practices (6%–11% across all included disruptions).

**Table 21**

*Circumstances for Students Not Taking Assessments at School*

| | Yes | | No | | Unknown | | Not applicable | |
|---|---|---|---|---|---|---|---|---|
| Circumstance | *n* | % | *n* | % | *n* | % | *n* | % |
| Student used different accessibility supports when testing remotely than at school | 121 | 11.5 | 361 | 34.3 | 63 | 6.0 | 506 | 48.2 |
| Student experienced technology difficulties during assessments taken remotely | 81 | 7.7 | 424 | 40.3 | 51 | 4.9 | 495 | 47.1 |
| Student had to respond in a less preferred response mode because of remote arrangements | 112 | 10.7 | 377 | 35.9 | 58 | 5.5 | 504 | 48.0 |
| Someone other than the teacher administered the assessments remotely | 65 | 6.2 | 435 | 41.4 | 47 | 4.5 | 504 | 47.9 |

Teachers responded to survey items about students' experience with testlets, which were also included on the 2017–2018 and 2018–2019 surveys. Teacher responses indicated that student experience with the testlets was similar across years, as shown in Table 22. The effect sizes for the differences in the percentages shown in Table 22 were all very small (less than .10)

**Table 22**

*Trends in Teacher Perceptions of Assessment Content and Student Experience*

| Teacher perception | 2017–2018 | | 2018–2019 | | 2020–2021 | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| The student responded to items to the best of their knowledge and ability | 49,772 | 89.5 | 47,384 | 89.6 | 33,977 | 89.4 |
| The student was able to respond regardless of disability, behavior, or health concerns | 47,176 | 84.6 | 44,916 | 84.9 | 32,624 | 86.5 |
| The student had access to all supports necessary to participate | 52,664 | 94.4 | 49,981 | 94.5 | 34,858 | 91.0 |

*Note*. Counts and percentages represent the teachers who responded *Agree* or *Strongly Agree*.

Table 23 shows the percentage of teachers who responded that most or all of the administered testlets matched instruction from 2017–2018 to 2020–2021. The percentage of teachers reporting that most or all testlets matched instruction was similar across all years. However, there was a slight increase of approximately 4 percentage points in all subjects from 2017–2018 to 2018–2019 (Cohen's *h* ranging from 0.07 to 0.08), followed by a small decline of 1–4 percentage points from 2018–2019 to 2020–2021 (Cohen's *h* ranging from 0.01 to 0.09).

**Table 23**

*Trends in Teacher Perceptions of Match Between Assessment Content and Instruction*

| Subject | 2017–2018 | | 2018–2019 | | 2020–2021 | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| Reading | 37,404 | 69.6 | 37,418 | 73.1 | 25,684 | 69.9 |
| Mathematics | 31,032 | 58.5 | 31,519 | 62.0 | 22,222 | 61.3 |
| Science | 18,541 | 54.4 | 19,640 | 58.3 | 12,633 | 53.9 |

*Note*. Counts and percentages represent the teachers who responded that most or all testlets matched instruction.

Teachers' confidence in their ability to deliver DLM testlets was fairly stable across years (97% agreed or strongly agreed in 2017–2018, and 94% agreed or strongly agreed in 2020–2021), as was the extent to which they used manuals and/or the DLM Educator Resource Page materials (91% and 90%, respectively). Teachers' perceptions of the extent to which test administrator training prepared them for responsibilities of test administrator decreased slightly (from 91% agreement in 2017–2018 to 86% agreement in 2020–2021, Cohen's $h = -0.16$).

# 5. Performance on DLM Assessments

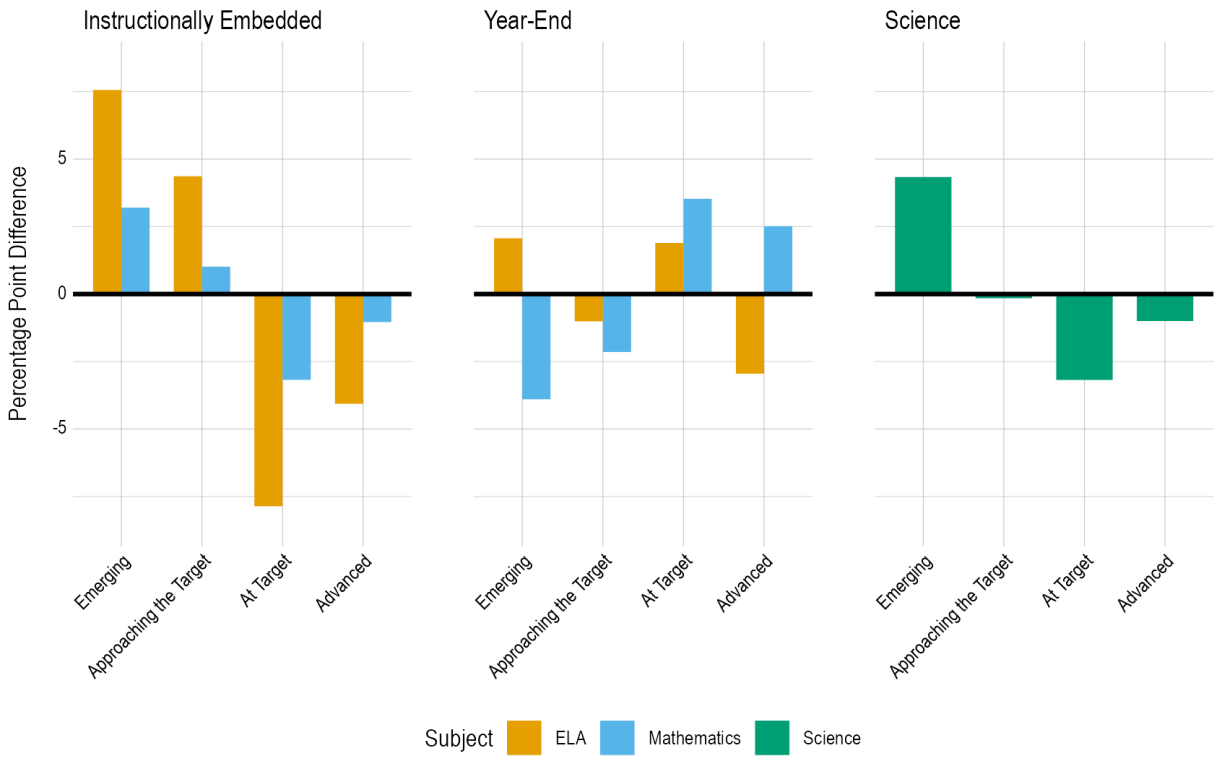
At the highest level of reporting, a performance level in each subject summarizes overall achievement on DLM assessments. For DLM assessments, there are four performance levels: *Emerging*, *Approaching the Target*, *At Target*, and *Advanced*. Performance levels are determined based on cut points applied to the total number of linkage levels for which a student demonstrates mastery in each subject. Mastery of linkage levels for each EE follows a linear hierarchy. That is, students are assigned mastery of all linkage levels below the highest level they demonstrate mastery of. For example, if a student demonstrates mastery of the Proximal Precursor linkage level, they are also assigned mastery of the Initial Precursor and Distal Precursor linkage levels. For a complete description of the DLM scoring method, see Chapter 5 of DLM Consortium (2017).

## 5.1. Overall Performance

To evaluate student performance in 2020–2021, we examined the difference in the percentage of students achieving at each performance level across years. Figure 3 shows the percentage point change from 2018–2019 to 2020–2021 for each assessment model and subject. Overall, students taking Instructionally Embedded ELA and mathematics and the DLM science assessments showed a decrease in performance from 2018–2019 to 2020–2021, with the percentage of students increasing at the Emerging and Approaching performance levels. Accordingly, the percentage of students achieving at the At Target and Advanced levels decreased from 2018–2019 to 2020–2021. The Year-End model had mixed findings. In ELA, the percentage of students increased at both the Emerging and At Target performance levels. In mathematics, there was an overall increase in performance. Across all models and subjects, the largest change was approximately 7.5 percentage points for Instructionally Embedded ELA (increase at the Emerging level and decrease at the At Target level).

**Figure 3**

*Overall Performance Level Changes From 2018–2019 to 2020–2021, by Model and Subject*



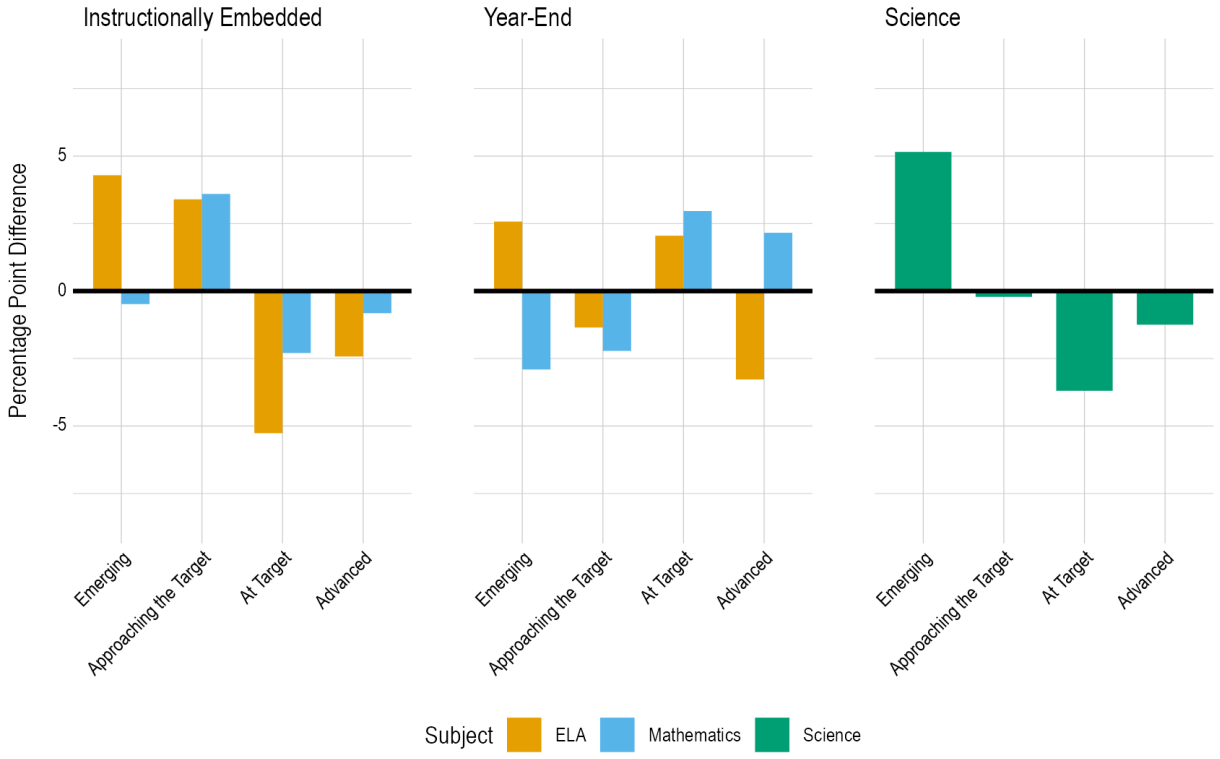
As described in the Participation section of this report, the demographic composition of the student population did show noticeable changes in 2020–2021, particularly within race, Hispanic ethnicity, and English learners. To evaluate whether the observed changes to the performance level distributions were a result of a change in performance rather than a change in the population, a propensity score matching approach was used, as described by Thompson and Hoover (2021). Using this method, performance level data from prior years (i.e., 2017–2018 and 2018–2019) were resampled to more closely resemble the sample of students that participated in 2020–2021. We estimated what the performance level distributions would have looked like in prior years if the student population looked like the 2020–2021 student population. In other words, we are evaluating what prior year performance distributions would look like with the student educational experiences of 2018–2019 but with the sample characteristics of 2020–2021. By controlling for the sample characteristics, any differences observed in the performance level distributions can be attributed to other factors, such as the students' educational experiences or changes to the assessment administration.

The propensity score matching algorithm used for comparing DLM performance levels preserves the original sample sizes of the two groups. Specifically for this analysis, the propensity score represents the probability that a student was a part of the 2020–2021 population, given their baseline characteristics. The 2018–2019 student data were resampled with replacement, weighted by the propensity score. Thus, the adjusted 2018–2019 sample looks more like the 2020–2021 sample, while preserving the original 2018–2019 sample size (see Table 3 for the sample sizes used for the propensity score models). However, it should be noted that there may be important covariates that are not collected, and are therefore not included in the propensity score model (e.g., urban/rural districts). These missing covariates may impact the overall effectiveness of the propensity score models. For details on the propensity score matching method, see Thompson and Hoover (2021).

Figure 4 shows the changes in the performance level distributions after adjusting for demographic changes from 2018–2019 to 2020–2021. Figure 4 demonstrates the same general patterns as Figure 3. However, the magnitude of many of the differences decreased (e.g., the Emerging level for Instructionally Embedded ELA decreased from ~7.5 in Figure 3 to ~4.0 in Figure 4). Thus, some of the observed differences in the raw performance levels changes can be attributed to changes in the student population. However, there are still notable changes after accounting for demographic changes. This suggests that there were factors other than demographic composition present in 2020–2021 that impacted student performance.

**Figure 4**

*Adjusted Performance Level Changes From 2018–2019 to 2020–2021, by Model and Subject*



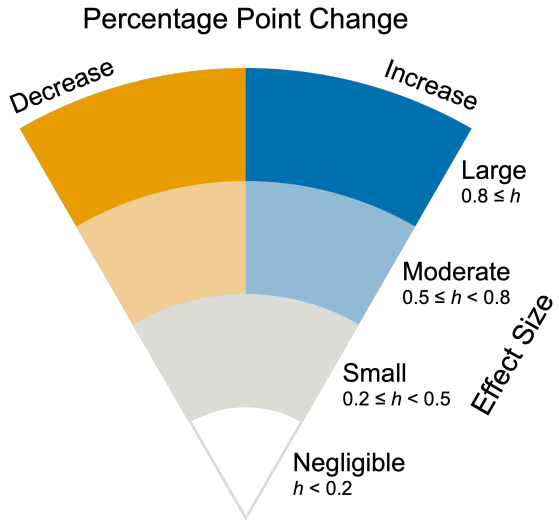

## 5.2. Model and Subject Results

The following sections provide more detailed performance results for Year-End ELA and mathematics, Instructionally Embedded ELA and mathematics, and science, respectively. For each assessment, we examine changes in the performance level distributions for each grade and subject, as well as for demographic subgroups. For all of these comparisons, we use the adjusted sample achieved through propensity score matching, due to the population shifts described in the Participation section of this report. Due to the number of comparisons, changes are highlighted according to their effect size. The shading methodology is based on value-suppressing color palettes (Correll et al., 2018), and is described in detail by Thompson and Hoover (2021). The cells in tables reporting results are shaded according to the legend in Figure 5. The color of the shading corresponds to the direction of the change, and the intensity of the color corresponds to the effect size of the change, where the size of the effect is defined using Cohen's *h* (Cohen, 1988). Percentage point changes with a negligible effect receive no shading. Changes corresponding to a small effect size receive gray shading, regardless of the direction of the change. Finally, for moderate and large effects, decreases are shaded in orange and increases shaded in blue. Moderate effect sizes are shaded with desaturated colors to differentiate moderate from large effects.

In some instances, performance level changes that are relatively large in magnitude may receive a small or negligible effect size. This most often happens when the sample sizes are small. When there are small samples, it is easier to see large changes in the percentage of students at each performance level. Thus, given that we expect larger changes with smaller samples, the magnitude needed for a change to be identified as moderate or large increases. That is, with smaller samples, we need stronger evidence (i.e., larger changes) to conclude that a meaningful change has occurred.

**Figure 5**

*Shading Palette for Classifying Performance Level Changes*




Because DLM performance levels are determined based on mastery of individual linkage levels for assessed EEs, we also examine scoring at the linkage level. This includes on which linkage levels students had the opportunity to demonstrate their knowledge and also performance on assessed linkage levels. If students tested at lower linkage levels, and/or if they mastered linkage levels less often, they would have fewer opportunities to demonstrate mastery (i.e., fewer total linkage levels mastered), potentially resulting in a lower performance level.

Linkage level performance is evaluated using *p*-values for items measuring each linkage level and the rate at which students mastered the linkage levels on which they were assessed. If students were assessed on lower linkage levels in 2020–2021, and the content of those linkage levels was too easy, we would expect *p*-values and mastery rates to increase. Conversely, if *p*-values and mastery rates are consistent with prior years, that would indicate that the linkage levels were appropriately difficult as intended, even though students were testing on lower linkage levels.

## *5.2.1. Instructionally Embedded ELA and Mathematics*

### 5.2.1.1. Performance Level Results

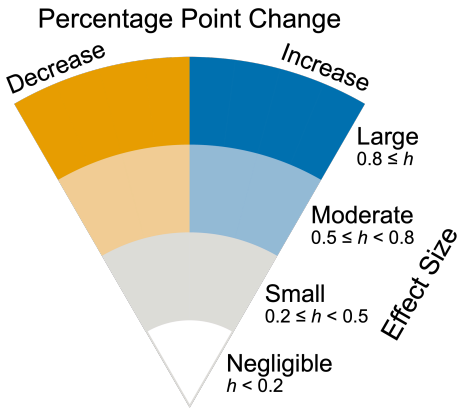
As described above, the Instructionally Embedded ELA and mathematics assessments had administration changes in addition to the potential impact due to the COVID-19 pandemic. Table 24 shows the percentage point change in performance level distributions by grade for Instructionally Embedded ELA and mathematics from 2018–2019 to 2020–2021, using the propensity-score adjusted sample. For example, the percentage of students in the Grade 3 Emerging performance level for ELA increased by 1.2 percentage points from 2018–2019 to 2020–2021. In total, there were five changes identified as non-negligible, all of which were in ELA and which were identified as *small* effect sizes. Overall, we observed a decrease in performance, with larger decreases in ELA.

We also examined performance level changes by demographic subgroup. Table 25 and Table 26 show the percentage point change in performance level distributions by grade for Instructionally Embedded ELA and mathematics, respectively, from 2017–2018 to 2018–2019 and from 2018–2019 to 2020–2021, after using the propensity score approach to match the 2017–2018 and 2018–2019 data to 2020–2021. The demographic subgroups showed stability, with only three *small* changes in ELA and no non-negligible changes in mathematics.

**Table 24**

*Instructionally Embedded Performance Level Changes, 2018–2019 to 2020–2021, by Grade*

| Performance level | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| English Language Arts | | | | | | | | | |
| Emerging | 1.2 | 1.8 | 2.1 | 7.3 | -0.2 | 2.9 | 14.2 | — | 4.8 |
| Approaching the Target | 0.5 | 4.8 | 6.3 | 3.4 | 12.0 | 5.4 | -0.6 | — | -0.6 |
| At Target | -0.8 | -4.0 | -5.2 | -7.0 | -7.5 | -6.8 | -9.3 | — | -2.8 |
| Advanced | -0.9 | -2.5 | -3.3 | -3.7 | -4.4 | -1.6 | -4.3 | — | -1.4 |
| *At Target/Advanced* | *-1.7* | *-6.6* | *-8.5* | *-10.7* | *-11.9* | *-8.3* | *-13.7* | — | *-4.2* |
| Mathematics | | | | | | | | | |
| Emerging | 4.1 | 1.0 | 2.9 | 0.0 | -1.2 | -0.6 | -1.5 | -2.3 | -3.5 |
| Approaching the Target | -1.8 | 2.4 | 1.3 | 4.8 | 3.2 | 3.3 | 6.8 | 4.3 | 8.0 |
| At Target | -0.5 | -3.1 | -2.7 | -3.8 | -0.3 | -1.7 | -2.6 | -0.4 | -2.2 |
| Advanced | -1.8 | -0.3 | -1.5 | -1.0 | -1.8 | -0.9 | -2.7 | -1.6 | -2.3 |
| *At Target/Advanced* | *-2.3* | *-3.4* | *-4.2* | *-4.8* | *-2.1* | *-2.6* | *-5.3* | *-2.0* | *-4.5* |




*Note*. Changes are based on adjusted samples from the propensity score matching algorithm. Due to changes in assessment administration changes should be interpreted with caution.

**Table 25**

*Instructionally Embedded ELA Performance Level Changes, 2017–2018 to 2020–2021, by Subgroup*

| Subgroup | Emerging | | Approaching the Target | | At Target | | Advanced | | At Target + Advanced | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 |
| Gender | | | | | | | | | | |
| Female | -1.2 | 4.8 | 0.7 | 3.1 | 1.3 | -4.2 | -0.8 | -3.7 | 0.5 | -7.9 |
| Male | 0.6 | 4.0 | -0.4 | 4.3 | -0.5 | -5.9 | 0.4 | -2.3 | -0.2 | -8.3 |
| Race | | | | | | | | | | |
| African American | 0.7 | 6.2 | 0.4 | 3.5 | -0.2 | -7.0 | -0.9 | -2.7 | -1.0 | -9.7 |
| Alaska Native | -26.3 | 27.0 | 29.0 | -37.2 | -2.7 | 7.2 | 0.0 | 3.0 | -2.7 | 10.2 |
| American Indian | -9.2 | -5.9 | 2.5 | 14.7 | 9.0 | -4.5 | -2.3 | -4.4 | 6.7 | -8.8 |
| Asian | 1.9 | 6.8 | -5.7 | 0.8 | 4.2 | -8.0 | -0.3 | 0.5 | 3.9 | -7.6 |
| Native Hawaiian or Pacific Islander | 17.4 | -1.7 | -1.4 | 7.8 | -1.4 | -6.0 | -14.5 | -0.1 | -15.9 | -6.1 |
| Two or more races | 4.6 | 0.7 | -4.5 | 5.9 | 0.1 | -3.7 | -0.2 | -2.9 | -0.1 | -6.5 |
| White | -0.4 | 4.3 | 0.3 | 3.8 | -0.1 | -5.1 | 0.3 | -2.9 | 0.1 | -8.0 |
| Hispanic ethnicity | | | | | | | | | | |
| Hispanic | -1.0 | 7.2 | 1.4 | 2.8 | -1.7 | -5.9 | 1.4 | -4.1 | -0.4 | -10.0 |
| Non-Hispanic | 0.0 | 3.9 | -0.2 | 4.0 | 0.4 | -5.3 | -0.2 | -2.6 | 0.1 | -7.9 |
| English learning (EL) participation | | | | | | | | | | |
| EL monitored or eligible | 4.7 | 8.9 | 1.7 | -1.6 | -9.1 | -5.5 | 2.8 | -1.7 | -6.3 | -7.3 |
| Not EL monitored or eligible | -0.3 | 4.0 | -0.1 | 4.2 | 0.6 | -5.3 | -0.2 | -2.9 | 0.4 | -8.2 |

*Note*. Highlighting indicates non-negligible effect size changes. Changes are based on adjusted samples from the propensity score matching algorithm. Due to changes in assessment administration changes should be interpreted with caution.

**Table 26**

*Instructionally Embedded Mathematics Performance Level Changes, 2017–2018 to 2020–2021, by Subgroup*

| Subgroup | Emerging | | Approaching the Target | | At Target | | Advanced | | At Target + Advanced | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 |
| Gender | | | | | | | | | | |
| Female | -0.2 | -0.1 | -0.5 | 2.8 | 0.1 | -1.8 | 0.6 | -0.9 | 0.7 | -2.6 |
| Male | -2.2 | 0.4 | 0.5 | 3.2 | 0.7 | -1.9 | 1.0 | -1.7 | 1.7 | -3.6 |
| Race | | | | | | | | | | |
| African American | 1.7 | 2.9 | -2.6 | 0.5 | 0.4 | -2.7 | 0.5 | -0.6 | 0.8 | -3.4 |
| Alaska Native | 7.9 | -14.4 | -5.5 | 8.3 | -2.4 | 4.5 | 0.0 | 1.5 | -2.4 | 6.1 |
| American Indian | -0.9 | 3.8 | 0.1 | -0.8 | 2.9 | -3.8 | -2.1 | 0.8 | 0.8 | -3.1 |
| Asian | -4.2 | 3.6 | 2.0 | -2.2 | 2.8 | -0.7 | -0.6 | -0.7 | 2.2 | -1.4 |
| Native Hawaiian or Pacific Islander | 1.8 | 12.3 | 3.0 | -18.5 | -6.3 | 4.9 | 1.5 | 1.3 | -4.8 | 6.2 |
| Two or more races | 2.8 | -0.8 | -3.7 | 3.4 | 0.8 | -2.8 | 0.2 | 0.2 | 0.9 | -2.6 |
| White | -2.4 | -0.4 | 0.8 | 3.9 | 0.4 | -1.7 | 1.2 | -1.8 | 1.6 | -3.4 |
| Hispanic ethnicity | | | | | | | | | | |
| Hispanic | -0.2 | 3.2 | -2.3 | 1.0 | 2.1 | -2.2 | 0.5 | -2.0 | 2.6 | -4.2 |
| Non-Hispanic | -1.7 | -0.2 | 0.4 | 3.3 | 0.3 | -1.8 | 0.9 | -1.3 | 1.2 | -3.1 |
| English learning (EL) participation | | | | | | | | | | |
| EL monitored or eligible | -1.2 | 4.5 | -0.8 | 1.1 | 1.3 | -2.7 | 0.7 | -2.9 | 2.1 | -5.7 |
| Not EL monitored or eligible | -1.5 | 0.0 | 0.2 | 3.2 | 0.4 | -1.8 | 0.9 | -1.3 | 1.3 | -3.1 |

*Note*. Changes are based on adjusted samples from the propensity score matching algorithm. Due to changes in assessment administration changes should be interpreted with caution.

#### 5.2.1.2. Linkage Level Results

Because teachers choose EEs and linkage levels in the Instructionally Embedded model, their choices may impact overall performance on the assessment. Table 27 shows the percentage of testlets administered at each linkage level by subject across the last three administration years. In 2017–2018 and 2018–2019, the linkage level was determined by the teacher during the fall window and assigned by the system for the subset of EEs administered in the spring window. In 2020–2021, teachers determined the linkage level in both windows. In both subjects, there was a noticeable increase in the percentage of testlets administered at the Initial Precursor and Distal Precursor linkage levels in 2020–2021. More research is needed to determine whether teacher selection was consistent with prior years, how system assignment may have affected 2017–2018 and 2018–2019 distributions, and how population shifts over time (e.g., as states work toward the 1% threshold for AA-AAS participation) may have impacted the 2020–2021 results.

**Table 27**

*Percentage of Instructionally Embedded Testlets Administered at Each Linkage Level*

| Linkage level | 2017–2018 (%) | 2018–2019 (%) | 2020–2021 (%) |
|---|---|---|---|
| English Language Arts | | | |
| Initial Precursor | 24.8 | 25.1 | 34.8 |
| Distal Precursor | 26.3 | 26.6 | 35.1 |
| Proximal Precursor | 20.8 | 20.5 | 14.9 |
| Target | 15.2 | 15.0 | 8.8 |
| Successor | 12.9 | 12.9 | 6.4 |
| Mathematics | | | |
| Initial Precursor | 31.5 | 29.9 | 39.2 |
| Distal Precursor | 29.8 | 29.7 | 33.7 |
| Proximal Precursor | 25.5 | 26.7 | 20.5 |
| Target | 9.7 | 9.6 | 5.6 |
| Successor | 3.6 | 4.1 | 0.9 |

We also examined how the teacher-selected linkage level compared to the system-recommended linkage level as a source of evidence for whether teacher selection potentially impacted overall assessment results. Table 28 shows the percentage of testlets administered at the recommended linkage level or adjusted above or below the system recommendation. For both fall and spring of 2020–2021, teachers chose the recommended linkage level much less often than in 2017–2018 or 2018–2019 and were more likely to adjust down than up. This pattern likely helps explain the results in Table 27 and the prevalence of the Initial Precursor and Distal Precursor linkage levels being tested. However, a higher rate of adjustment does not appear to be a result of the COVID-19 pandemic. Also included in Table 28 are data from fall of 2019–2020, preceding the onset of pandemic-related school closures. Fall of 2019 was the first assessment window with the new Instruction and Assessment Planner tool for selecting assessment content. Thus, present evidence suggests that there are other factors driving teachers to deviate from the recommended linkage level, unrelated to COVID-19.

**Table 28**

*Teacher Adjustment of Recommended Linkage Level*

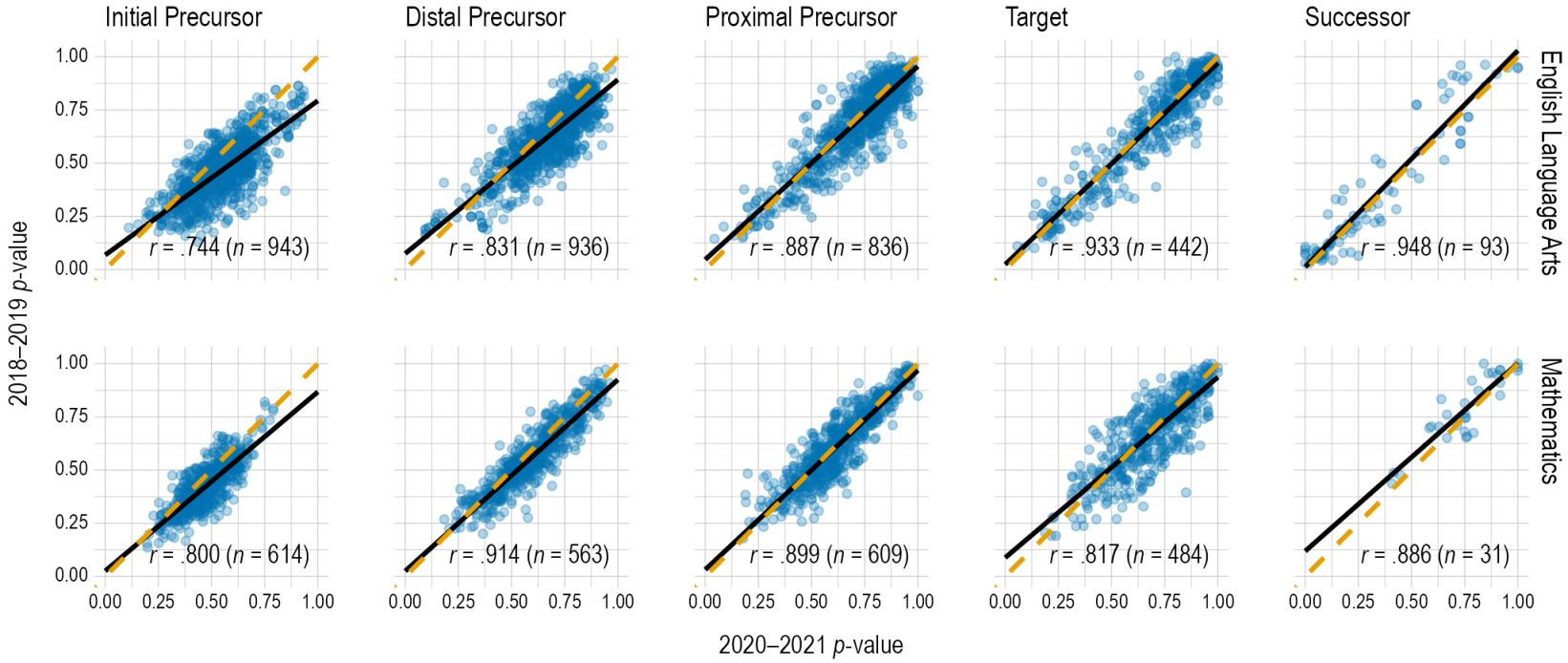
| Linkage level selection | 2017–2018 Fall (%) | 2018–2019 Fall (%) | 2019–2020 Fall (%) | 2020–2021 Fall (%) | 2020–2021 Spring (%) |
|---|---|---|---|---|---|
| English Language Arts | | | | | |
| Adjust down | 12.1 | 10.6 | 33.1 | 31.5 | 26.1 |
| Accept recommendation | 83.8 | 86.1 | 59.5 | 60.7 | 62.2 |
| Adjust up | 4.1 | 3.3 | 7.4 | 7.8 | 11.7 |
| Mathematics | | | | | |
| Adjust down | 15.7 | 13.1 | 37.3 | 35.9 | 30.4 |
| Accept recommendation | 80.5 | 83.7 | 55.4 | 56.9 | 59.9 |
| Adjust up | 3.8 | 3.2 | 7.3 | 7.2 | 9.8 |

*Note*. 2017–2018 and 2018–2019 predate the release of the Instruction and Assessment Planner tool for assigning testlets. Fall recommended level is based on student complexity band as determined by the First Contact survey. Spring 2021 recommended level is based on fall 2020 performance for the EE and linkage level, if tested. If not previously tested, the recommended level is based on the complexity band.

Because of the acknowledged impact of the COVID-19 pandemic on student learning experiences, we explored whether teachers might be selecting a lower level to account for potential learning loss. We evaluated *p*-values for operational items over time to determine if items appeared to be more challenging for students during the COVID-impacted 2020–2021 administration. Figure 6 shows the *p*-values for operational items administered in both 2018–2019 and 2020–2021. Here we see a shift toward higher *p*-values in 2020–2021, compared to 2018–2019, especially at the Initial Precursor and Distal Precursor linkage levels. This means that students are answering those items correctly more often now than in previous years. Thus, the evidence suggests that teachers may be selecting linkage levels lower than what their students could achieve. This is supported by the rate of linkage level mastery. In 2020–2021 Instructionally Embedded ELA and mathematics, students mastered the linkage level they were assessed on 90.3% of the time. In 2017–2018 and 2018–2019, the mastery rates were 86.3% and 86.8%, respectively. Thus, the rate of mastery increased by 3.5–4.0 percentage points in 2020–2021. However, we do not know if instruction prepared students to demonstrate mastery on a higher linkage level for the EE.

**Figure 6**

*Operational* p*-values for Instructionally Embedded Assessment Items*




*Note*. The orange dashed line represents perfect agreement, and the solid black line is the linear best fit.

In totality, the analyses from the Instructionally Embedded model indicate that there was a decrease in performance from 2018–2019 to 2020–2021 and that this decrease was likely driven at least in part by teachers' selection of lower linkage levels. However, it appears that the selection of lower linkage levels pre-dates the COVID-19 pandemic, and therefore, it is unclear how much of the decrease is due to true changes in performance, shifts in the population, teachers' use of the new Instruction and Assessment Planner tool, or teacher expectations for what students know and can do.

### *5.2.2. Year-End ELA and Mathematics*

#### 5.2.2.1. Performance Level Results

As described above, the Year-End ELA and mathematics assessments saw significant changes in 2019–2020, unrelated to the COVID-19 pandemic. Because the full cut point adjustment process could not be completed due to the COVID-19 pandemic, preliminary cut points were set based on the expected distribution of student performance based on 2018–2019 data. Given the changes to the assessment, direct comparisons across years should be interpreted with extreme caution.

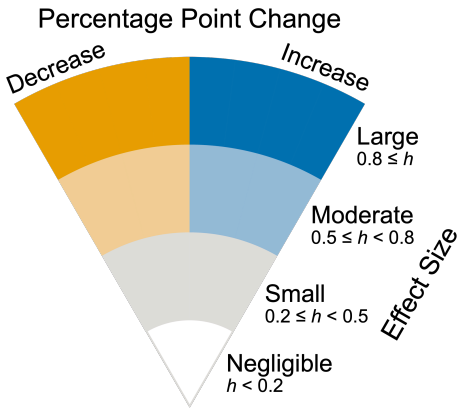
Table 29 shows the percentage point change in performance level distributions by grade for Year-End ELA and mathematics using the propensity-score adjusted sample. For example, the percentage of students in the Grade 3 Emerging performance level for ELA decreased by 3.9 percentage points from 2018–2019 to 2020–2021. In total, there were 12 changes identified as non-negligible, 10 of which were identified as having a *small* effect size, and *2* as having a moderate effect size. In mathematics, we generally observed an increase in performance, with the percentage of students at the At Target or Advanced levels increasing is almost all grades. Year-End ELA assessments showed more mixed findings, with some grades showing an increase in performance and other grades showing a decrease.

We also examined performance level changes by demographic subgroup. Table 30 and 31 show the percentage point change in performance level distributions by grade for Year-End ELA and mathematics, respectively, from 2017–2018 to 2018–2019 and from 2018–2019 to 2020–2021, after using the propensity score approach to match the 2017–2018 and 2018–2019 data to 2020–2021. Overall, the subgroup distributions were stable, with no non-negligible effects size changes identified.

**Table 29**

*Year-End Performance Level Changes, 2018–2019 to 2020–2021, by Grade*

| Performance level | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| English Language Arts | | | | | | | | | |
| Emerging | -2.8 | 7.8 | 3.1 | -5.6 | 9.6 | 4.6 | -3.3 | -8.6 | -1.0 |
| Approaching the Target | -6.2 | -6.2 | -9.9 | -2.6 | -3.5 | 4.1 | 2.4 | 5.3 | -0.5 |
| At Target | 8.3 | -1.2 | 6.4 | 8.0 | -2.8 | -0.4 | 5.0 | 5.1 | -0.1 |
| Advanced | 0.7 | -0.4 | 0.5 | 0.2 | -3.3 | -8.4 | -4.1 | -1.7 | 1.6 |
| *At Target/Advanced* | *9.0* | *-1.6* | *6.8* | *8.2* | *-6.1* | *-8.8* | *0.9* | *3.3* | *1.5* |
| Mathematics | | | | | | | | | |
| Emerging | 2.5 | -4.3 | -16.9 | 1.5 | -0.3 | -2.0 | 4.6 | -7.0 | -17.0 |
| Approaching the Target | -0.9 | -5.7 | 7.4 | -4.1 | -4.7 | 5.1 | -14.3 | 2.4 | -7.1 |
| At Target | -3.2 | 0.1 | 6.8 | 1.3 | 4.6 | -3.9 | 11.0 | 2.9 | 20.8 |
| Advanced | 1.5 | 9.9 | 2.7 | 1.3 | 0.4 | 0.8 | -1.3 | 1.6 | 3.4 |
| *At Target/Advanced* | *-1.7* | *10.1* | *9.5* | *2.6* | *5.0* | *-3.1* | *9.7* | *4.5* | *24.2* |




*Note*. Changes are based on adjusted samples from the propensity score matching algorithm. Due to changes in assessment administration changes should be interpreted with caution.

**Table 30**

*Year-End ELA Performance Level Changes, 2017–2018 to 2020–2021, by Subgroup*

| Subgroup | Emerging | | Approaching the Target | | At Target | | Advanced | | At Target + Advanced | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 |
| Gender | | | | | | | | | | |
| Female | 0.4 | 2.1 | 2.7 | -3.7 | -1.5 | 3.4 | -1.6 | -1.8 | -3.1 | 1.6 |
| Male | 0.8 | 2.4 | 2.4 | -3.6 | -1.4 | 2.8 | -1.8 | -1.6 | -3.2 | 1.2 |
| Race | | | | | | | | | | |
| African American | 0.9 | 6.1 | 3.8 | -5.7 | -2.1 | 1.7 | -2.5 | -2.2 | -4.7 | -0.4 |
| Alaska Native | 2.7 | 5.1 | -1.8 | -8.4 | -0.3 | 2.1 | -0.6 | 1.1 | -0.9 | 3.2 |
| American Indian | 1.4 | 6.4 | 3.6 | -1.4 | -3.8 | -1.4 | -1.2 | -3.6 | -5.0 | -5.0 |
| Asian | 1.1 | 0.7 | 1.2 | -2.3 | 0.0 | 1.9 | -2.3 | -0.3 | -2.3 | 1.6 |
| Native Hawaiian or Pacific Islander | 8.9 | 15.3 | -2.0 | -6.0 | 0.1 | -2.2 | -7.0 | -7.1 | -6.9 | -9.3 |
| Two or more races | -1.0 | 2.5 | 4.3 | -4.4 | -2.1 | 2.6 | -1.2 | -0.7 | -3.3 | 1.9 |
| White | 0.5 | 0.9 | 1.8 | -3.0 | -1.0 | 3.8 | -1.4 | -1.7 | -2.3 | 2.1 |
| Hispanic ethnicity | | | | | | | | | | |
| Hispanic | 0.9 | 4.0 | 2.7 | -3.2 | -1.9 | 1.5 | -1.7 | -2.3 | -3.6 | -0.8 |
| Non-Hispanic | 0.6 | 1.9 | 2.4 | -3.8 | -1.3 | 3.4 | -1.7 | -1.5 | -3.0 | 1.9 |
| English learning (EL) participation | | | | | | | | | | |
| EL monitored or eligible | 1.0 | 5.9 | 1.9 | -5.7 | -0.8 | 1.3 | -2.1 | -1.6 | -2.9 | -0.2 |
| Not EL monitored or eligible | 0.6 | 2.1 | 2.5 | -3.5 | -1.5 | 3.1 | -1.7 | -1.7 | -3.1 | 1.4 |

*Note*. Changes are based on adjusted samples from the propensity score matching algorithm. Due to changes in assessment administration changes should be interpreted with caution.

**Table 31**

*Year-End Mathematics Performance Level Changes, 2017–2018 to 2020–2021, by Subgroup*

| Subgroup | Emerging | | Approaching the Target | | At Target | | Advanced | | At Target + Advanced | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 |
| Gender | | | | | | | | | | |
| Female | 2.8 | -4.6 | -0.8 | -1.5 | -1.1 | 4.1 | -0.9 | 1.9 | -2.0 | 6.1 |
| Male | 1.7 | -4.0 | 0.2 | -2.4 | -1.0 | 3.2 | -0.8 | 3.1 | -1.8 | 6.3 |
| Race | | | | | | | | | | |
| African American | 2.9 | -2.2 | -0.3 | -2.5 | -1.7 | 2.5 | -0.9 | 2.2 | -2.6 | 4.7 |
| Alaska Native | -1.7 | -3.3 | 0.1 | -5.3 | 0.9 | 4.7 | 0.7 | 3.9 | 1.6 | 8.6 |
| American Indian | 4.5 | 1.8 | -2.0 | -1.7 | -1.0 | 0.0 | -1.6 | -0.1 | -2.6 | -0.1 |
| Asian | 0.6 | -2.2 | 1.5 | -2.9 | -1.9 | 2.6 | -0.2 | 2.6 | -2.1 | 5.1 |
| Native Hawaiian or Pacific Islander | -1.0 | 18.0 | 0.7 | -10.1 | 4.5 | -8.6 | -4.2 | 0.7 | 0.3 | -7.9 |
| Two or more races | 2.0 | -6.0 | -0.1 | -0.2 | -0.3 | 3.6 | -1.7 | 2.6 | -1.9 | 6.2 |
| White | 1.5 | -5.2 | -0.1 | -2.2 | -0.8 | 4.2 | -0.7 | 3.1 | -1.5 | 7.3 |
| Hispanic ethnicity | | | | | | | | | | |
| Hispanic | 3.2 | -1.6 | -0.1 | -1.5 | -1.3 | 1.3 | -1.8 | 1.8 | -3.1 | 3.1 |
| Non-Hispanic | 1.8 | -4.8 | -0.1 | -2.2 | -1.0 | 4.1 | -0.6 | 3.0 | -1.6 | 7.1 |
| English learning (EL) participation | | | | | | | | | | |
| EL monitored or eligible | -0.9 | 3.8 | 0.0 | -3.9 | 0.3 | -0.4 | 0.5 | 0.4 | 0.8 | 0.0 |
| Not EL monitored or eligible | 2.2 | -4.7 | -0.2 | -2.0 | -1.1 | 3.8 | -1.0 | 2.9 | -2.1 | 6.6 |

*Note*. Changes are based on adjusted samples from the propensity score matching algorithm. Due to changes in assessment administration changes should be interpreted with caution.

#### 5.2.2.2. Linkage Level Results

The Year-End model uses adaptive routing to assign the testlet linkage level. When students respond to <35% of items correctly, they are routed down to a lower linkage level on the next testlet. Conversely, when students respond to >80% of items correctly, they are routed up a linkage level on the next testlet. Any value in between results in the student staying at the same linkage level on the subsequent testlet. For a complete description of the adaptive routing, see Chapter 4 of DLM Consortium (2016b). In 2020–2021, there were changes in the percentage of testlets administered at each linkage level. Table 32 indicates that Year-End model students shifted toward testing on the higher linkage levels more often in 2020–2021 than in the previous year, especially in mathematics.

**Table 32**

*Percentage of Year-End Testlets Administered at Each Linkage Level*

| Linkage level | 2017–2018 (%) | 2018–2019 (%) | 2020–2021 (%) |
|---|---|---|---|
| English Language Arts | | | |
| Initial Precursor | 22.8 | 26.4 | 25.5 |
| Distal Precursor | 20.8 | 24.2 | 23.8 |
| Proximal Precursor | 21.7 | 19.7 | 17.1 |
| Target | 17.6 | 15.2 | 15.1 |
| Successor | 17.1 | 14.5 | 18.5 |
| Mathematics | | | |
| Initial Precursor | 40.6 | 41.4 | 32.4 |
| Distal Precursor | 26.3 | 26.5 | 24.2 |
| Proximal Precursor | 18.6 | 18.5 | 21.1 |
| Target | 9.4 | 9.0 | 12.5 |
| Successor | 5.0 | 4.6 | 9.8 |

The shift toward higher linkage levels could be attributed to changes in the population or the change in the operational item pool impacting adaptive routing decisions between testlets. The student population had fewer students from the Foundational complexity band testing in 2020–2021, as described in the Participation section of this report. The Foundational students test primarily at Initial Precursor and Distal Precursor linkage levels. Therefore, a decrease in the prevalence of these students in the population could be associated with fewer testlets being administered at those linkage levels.

Alternatively, under the original blueprint, the higher number of EEs meant that a single testlet measured multiple EEs, often with only one or two items. To determine the linkage level of subsequent testlets, the percentage of items answered correctly for each EE was calculated, and the lowest percentage determined the routing. When an EE was measured by only one item, and that item was answered incorrectly, the lowest percentage correct would be 0% (i.e., 0 out of 1), and the student would route down to a lower linkage level on the next testlet, even if all other items on the testlet were answered correctly.

Under the revised blueprint, the reduced number of EEs means that each testlet measures only one EE with typically three to five items. Getting one item wrong on a single-EE testlet would result in a percentage correct between 66% (2 out of 3) and 80% (4 out of 5). This means that a single incorrect response will never result in a decision to route down and could potentially even result in routing up (i.e., 80% would route up). Thus, the change to single-EE testlets makes it easier for students to remain at their current linkage level, or route up. The result is that students will test on

higher linkage levels because they are routing down and testing on the lower linkage levels less often. However, because the routing decisions are based on a greater number of items, we can be more confident in those decisions.

Further, the data do not suggest that students were performing more poorly on the higher linkage levels. Due to the change in the item pool, it is not possible to compare *p*-values from 2018–2019 to 2020–2021 because the item pools were distinct (i.e., none of ELA and mathematics testlets administered to Year-End students in 2018–2019 were administered 2020–2021). However, we can evaluate how often students mastered the linkage levels they were assessed on. In 2020–2021, students demonstrated mastery of the linkage level they were assessed on 78% of the time. For comparison, this number was 86% in 2018–2019 and 85% in 2017–2018. Thus, there was a small decrease in 2020–2021; although, as mentioned previously, the mastery decisions in 2020–2021 are also more reliable due to the increased number of items, and students were generally assessed on higher linkage levels.

In summary, the distribution of linkage levels assessed and mastery rates indicate that performance may have increased in 2020–2021. However, the impact these changes have on performance levels and how that relates to impacts stemming from COVID-19 remain unclear due to the confounding changes to the Year-End ELA and mathematics assessments.

### *5.2.3. Science*

#### 5.2.3.1. Performance Level Results

As noted above, there were no changes to the science administration or blueprint from 2018–2019 to 2020–2021. Thus, the primary influences on performance level changes from 2018–2019 to 2020–2021 are the COVID-19 pandemic and population changes. Table 33 shows the percentage point change in performance level distributions by grade or course for science from 2018–2019 to 2020–2021 using the adjusted samples derived from the propensity score matching process. As was shown in Figures 3 and 4, there were small but consistent decreases in the distribution of performance levels across grades. However, no changes had a non-negligible effect size.

We also examined performance level changes by demographic subgroup. Table 34 shows the percentage point change in performance level distributions by grade for science, from 2017–2018 to 2018–2019 and from 2018–2019 to 2020–2021, after using the propensity score approach to match the 2017–2018 and 2018–2019 data to 2020–2021. Overall, the subgroup distributions were stable, with no non-negligible effects size changes.

**Table 33**

*Science Performance Level Changes, 2018–2019 to 2020–2021, by Grade*

| Performance level | 3 | 4 | 5 | 6 | 7 | 8 | 9–12 | Biology |
|---|---|---|---|---|---|---|---|---|
| Emerging | 5.5 | 6.7 | 2.7 | 9.8 | 5.3 | 5.1 | 4.3 | -4.5 |
| Approaching the Target | -1.8 | 0.2 | -0.4 | 0.2 | -0.9 | -1.2 | 0.6 | -3.6 |
| At Target | -0.6 | -4.0 | -2.1 | -6.7 | -3.6 | -3.0 | -3.5 | 6.8 |
| Advanced | -3.1 | -3.0 | -0.2 | -3.3 | -0.8 | -0.9 | -1.3 | 1.3 |
| *At Target/Advanced* | *-3.7* | *-6.9* | *-2.3* | *-10.0* | *-4.4* | *-3.9* | *-4.9* | *8.2* |

*Note*. Changes are based on adjusted samples from the propensity score matching algorithm.

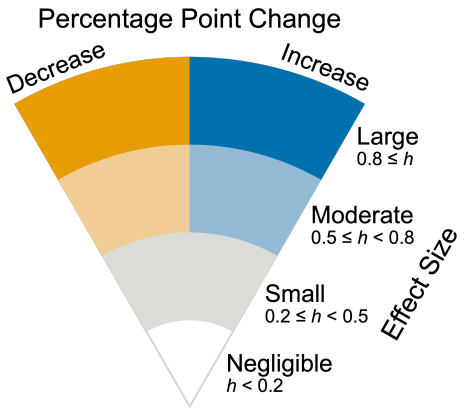


**Table 34**

*Science Performance Level Changes, 2017–2018 to 2020–2021, by Subgroup*

| Subgroup | Emerging | | Approaching the Target | | At Target | | Advanced | | At Target + Advanced | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 | 2018 to 2019 | 2019 to 2021 |
| Gender | | | | | | | | | | |
| Female | 0.7 | 4.2 | 0.7 | -0.3 | -1.6 | -2.8 | 0.2 | -1.1 | -1.4 | -3.9 |
| Male | 1.0 | 4.6 | -0.5 | -0.2 | -0.7 | -3.3 | 0.3 | -1.1 | -0.5 | -4.4 |
| Race | | | | | | | | | | |
| African American | 3.8 | 6.2 | -0.7 | -0.5 | -2.6 | -3.5 | -0.5 | -2.1 | -3.1 | -5.7 |
| Alaska Native | 5.6 | -11.0 | -14.6 | 7.1 | 7.8 | 2.0 | 1.2 | 1.9 | 9.0 | 3.9 |
| American Indian | -1.1 | 9.8 | -3.8 | 1.4 | 2.9 | -6.9 | 2.0 | -4.4 | 4.9 | -11.2 |
| Asian | 0.1 | 2.1 | -1.9 | 0.8 | 1.4 | -2.4 | 0.4 | -0.5 | 1.8 | -2.9 |
| Native Hawaiian or Pacific Islander | 3.7 | 19.4 | -0.5 | 1.3 | -1.7 | -16.3 | -1.6 | -4.4 | -3.3 | -20.7 |
| Two or more races | -0.5 | 5.6 | 1.8 | -1.1 | -0.8 | -3.5 | -0.4 | -1.0 | -1.3 | -4.5 |
| White | 0.1 | 3.8 | 0.2 | -0.2 | -0.8 | -2.9 | 0.5 | -0.7 | -0.3 | -3.6 |
| Hispanic ethnicity | | | | | | | | | | |
| Hispanic | 1.4 | 7.0 | 0.1 | -1.1 | -1.9 | -4.1 | 0.4 | -1.8 | -1.5 | -6.0 |
| Non-Hispanic | 0.8 | 4.0 | -0.2 | -0.1 | -0.8 | -3.0 | 0.2 | -1.0 | -0.6 | -3.9 |
| English learning (EL) participation | | | | | | | | | | |
| EL monitored or eligible | 3.1 | 9.2 | -2.1 | 1.4 | -0.9 | -8.1 | -0.1 | -2.5 | -0.9 | -10.6 |
| Not EL monitored or eligible | 0.8 | 4.3 | 0.0 | -0.3 | -1.0 | -2.9 | 0.2 | -1.0 | -0.8 | -3.9 |

*Note*. Highlighting indicates non-negligible effect size changes. Changes are based on adjusted samples from the propensity score matching algorithm.

#### 5.2.3.2. Linkage Level Results

In science, as in the Year-End ELA and mathematics assessments, linkage levels are system-assigned through adaptive routing. Table 35 shows the percentage of testlets administered at each linkage level. In 2020–2021 there was a small increase in the percentage of testlets administered at the Initial and Precursor linkage levels. Thus, students were testing at slightly lower linkage levels than in previous administration years. Because linkage levels in science are determined by the adaptive routing algorithm, this means that students were answering fewer items correct in 2020–2021, leading to more instances of routing down to lower linkage levels or remaining at the same linkage level instead of routing up.

**Table 35**

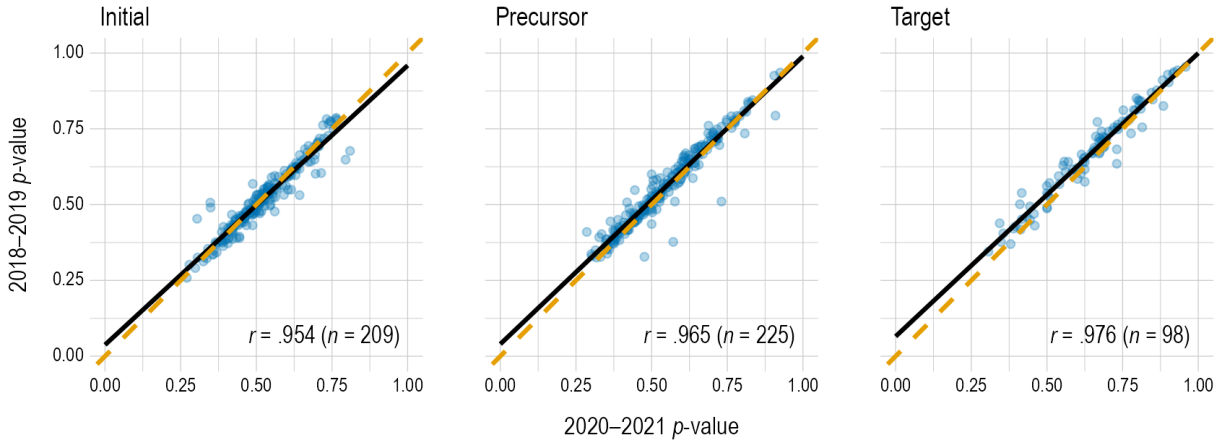
*Percentage of Science Testlets Administered at Each Linkage Level*

| Linkage level | 2017–2018 (%) | 2018–2019 (%) | 2020–2021 (%) |
|---|---|---|---|
| Initial | 35.2 | 36.2 | 38.1 |
| Precursor | 34.2 | 34.2 | 35.3 |
| Target | 30.7 | 29.5 | 26.7 |

Figure 7 compares performance on linkage levels by examining the *p*-values for operational assessment items in 2020–2021 compared to 2018–2019. In Figure 7, the dashed orange line indicates perfect association, and the solid black line is the linear best fit. For all three linkage levels, there was a very strong relationship, indicating the performance on items for each linkage level was nearly the same in both years. Thus, the evidence suggests that although students were testing on lower linkage levels, the content was appropriately difficult, as intended. That is, we did not see an increase in *p*-values, as we might if students were testing on linkage levels that were too easy. In fact, we actually observed a very small decrease in the *p*-values, indicating that items may have been slightly more difficult in 2020–2021. Linkage level mastery rates also provide evidence that the lower linkage levels were still appropriately difficult. In 2020–2021, students mastered the linkage level they were assessed on 86% of the time. This compares to 84% in 2017–2018 and 85% in 2018–2019. Thus, students were mastering the linkage levels assigned to them at similar rates, even though they were being assigned slightly lower linkage levels overall.

**Figure 7**

*Operational* p*-values for Science Assessment Items*




*Note*. The orange dashed line represents perfect agreement, and the solid black line is the linear best fit.

In total, the evidence from the DLM science assessment indicates that student performance slightly decreased in 2020–2021. Students answered questions correctly less often, resulting in students testing on lower linkages at a higher rate than in previous years. Additionally, students mastered the linkage levels they were assessed on at a similar rate to prior years. Because students tested on lower linkage levels, on average, but did not master those linkage levels at a higher rate, the number of total linkage levels mastered also decreased, likely resulting in the overall decrease in the performance level distributions.

# 6. Discussion

The purpose of this report was to analyze multiple available data sources and describe students' assessment participation, educational experiences, and performance during the COVID-19 pandemic in 2020–2021.

## 6.1. Summary of Findings

The rate of participation varied by state, from approximately half of students participating to near full participation relative to previous administration years. Rates were influenced by local policy and pandemic severity, and they may also be influenced by states working toward the 1% threshold on students participating in alternate assessments. Large districts had lower participation rates, suggesting urban districts were more heavily impacted by COVID-19. Across the consortium, we observed reduced rates of historically marginalized students participating in the assessment, as well as students classified to the lowest complexity level. More research is needed to determine whether these student groups with reduced assessment participation also had reduced access to instruction during the year.

Students' educational experiences also varied. Most students received in-school instruction for at least a part of the year and also experienced changes between in-person and remote learning. We observed some small subgroup differences in instructional conditions during the pandemic, but these may reflect several factors, including different levels of state participation, response bias in the survey data, and that students from historically marginalized populations often attend urban or city school districts that are widely acknowledged as being the most severely impacted by the pandemic.

Some of the survey results suggest that the pandemic may have had an impact on students' opportunity to learn, but these results may at least partially reflect differences in both the student population testing and those with survey responses in 2020–2021 compared to prior years. In all subjects, students were more likely to receive less than 5 hours of instruction and were less likely to receive more than 20 hours of instruction compared to prior academic years. In addition, the survey data suggest that breadth and depth of instruction declined slightly from 2018–2019 to 2020–2021. Yet, the portion of testlets that matched instruction remained relatively stable in 2017–2018, 2018–2019, and 2020–2021; and when looking only at students with matched survey data, the majority received the same amount or a greater amount of instruction in 2020–2021 compared to 2018–2019. Due to these mixed results, additional research is needed to better understand students' opportunity to learn during the pandemic.

Teachers also described a variety of factors impacting students' instruction during the pandemic, including difficulties with remote learning (e.g., technology issues, lack of access to materials, lack of student engagement and parental support), COVID-19 safety protocols and social distancing, students' mental and physical health, and family hardships. Teachers described benefits for students who were able to attend school in person.

According to special circumstances codes and teacher survey results, only a small group of students took assessments outside of school and/or were administered DLM assessments by someone other than their primary teacher. Most students taking assessments outside of school did not have difficulties with accessibility supports, technology, or response modes. During 2020–2021, ATLAS released a policy stating that DLM assessments must be administered in person, not virtually, by a qualified test administrator; however, in-school administration was not a requirement. Because of the teacher survey wording (i.e., use of "remote"), we were unable to disentangle whether remote administration made use of a trained, in-person test administrator. More research is needed to know whether misadministration occurred for students who completed assessments remotely.

Students who completed DLM assessments generally were able to complete all expected assessments. Blueprint coverage rates in states adopting the Year-End model and science remained stable, with nearly all students completing all expected assessments. Blueprint coverage increased for states adopting the Instructionally

Embedded model and may be attributable to the release of the Instruction and Assessment Planner tool for selection of assessment content. More research is needed to evaluate how the Planner tool impacted teachers' selection of testlets.

Performance on DLM assessments also varied and was likely impacted by assessment model changes that were unrelated to the COVID-19 pandemic. These changes complicated the evaluation of assessment performance across years. We observed a slight increase in performance on Year-End model ELA and mathematics assessments. Many factors likely contributed to this change, including shifts in the student population, the implementation of a revised assessment blueprint, changes to the operational item pool, the use of administrative cut points, and students testing at higher linkage levels, in addition to any effects stemming from the COVID-19 pandemic.

In contrast, we observed a decrease in performance on Instructionally Embedded ELA and mathematics assessments. Again, many factors likely contributed to this finding, including changes to the student population, the adjustment to two instructionally embedded administration windows, and teacher selection of lower linkage levels for instruction and assessment. However, the evidence suggests that the shift in linkage level selection was unrelated to the COVID-19 pandemic, as the change was first observed during the fall 2019 assessment window.

Decreases in performance were also observed for the science assessment. Because there were no changes to assessment administration in science, the results from the science assessment provided the clearest comparison for assessing the impact of the COVID-19 pandemic on student performance. Students were assessed on lower linkages in science, on average, when compared to prior years. Because linkage levels in science are system-assigned and the linkage levels appeared still be appropriately difficult, we can conclude that student performance did in fact decrease in 2020–2021. This suggests that, at least for the science assessment, the COVID-19 pandemic did have some effect on student achievement.

## 6.2. Future Directions

More research is needed to fully understand the effect of COVID-19 on students' participation, educational experience, and performance on DLM assessments. Although we use all data available to us, some data that might have been useful for analyses related to COVID-19 are not available, limiting the scope of possible analyses. For example, it is widely acknowledged that urban districts, which tend to have higher populations of students from historically marginalized groups, were more impacted by the pandemic than more rural districts. However, we do not collect data on the population density for individual schools or districts and therefore cannot make comparisons between urban and rural settings. Similarly, we do not collect information about a student's socioeconomic status. Thus, many comparisons that may be relevant to evaluating the effect of COVID-19 on DLM administration were not possible due to constraints in the data. Additionally, many findings showed mixed results (e.g., opportunity to learn) or were confounded by changes to the assessment administration (e.g., performance level distributions). These mixed and confounded findings also limited subsequent analyses. For example, we did not examine the relationship between opportunity to learn and overall performance due to the limitations of the individual analyses. Future work will continue to evaluate these aspects of the DLM assessments and be published for public consumption.

# 7. References

Centers for Disease Control and Prevention. (2020). Considerations for school closure. https://www.cdc.gov/coronavirus/2019-ncov/downloads/considerations-for-school-closure.pdf

Clark, A. K., Thompson, W. J., Kobrin, J., Kavitsky, E., & Karvonen, M. (2021). *The impact of COVID-19: Validity considerations and scoring and reporting in flexible scenarios* (Project Brief No. 21-01). University of Kansas, Accessible Teaching, Learning, and Assessment Systems. https://dynamiclearningmaps.org/sites/default/files/documents/publication/Project_Brief_Impact_of_COVID-19.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd). L. Erlbaum Associates.

Correll, M., Moritz, D., & Heer, J. (2018, April 21–26). *Value-suppressing uncertainty palettes* [Paper presentation], In *Proceedings of the 2018 CHI conference on human factors in computing systems*. CHI '18: CHI Conference on Human Factors in Computing Systems, Montreal, Quebec, Canada. 1–11. https://doi.org/10.1145/3173574.3174216.

Dynamic Learning Maps Consortium. (2016a). *2014–2015 Technical Manual—Integrated Model*. University of Kansas, Center for Educational Testing and Evaluation. https://dynamiclearningmaps.org/sites/default/files/documents/publication/Technical_Manual_IM_2014-15.pdf

Dynamic Learning Maps Consortium. (2016b). *2014–2015 Technical Manual—Year-End Model*. University of Kansas, Center for Educational Testing and Evaluation. https://dynamiclearningmaps.org/sites/default/files/documents/publication/Technical_Manual_YE_2014-15.pdf

Dynamic Learning Maps Consortium. (2017). *2015–2016 Technical Manual Update—Integrated Model*. University of Kansas, Center for Educational Testing and Evaluation. https://dynamiclearningmaps.org/sites/default/files/documents/publication/DLM_Technical_Manual_IM_2015-16.pdf

Dynamic Learning Maps Consortium. (2020a). *2019–2020 Technical Manual Update—Instructionally Embedded Model*. University of Kansas, Accessible Teaching, Learning, and Assessment Systems. https://2020-ie-techmanual.dynamiclearningmaps.org/

Dynamic Learning Maps Consortium. (2020b). *2019–2020 Technical Manual Update—Year-End Model*. University of Kansas, Accessible Teaching, Learning, and Assessment Systems. https://2020-ye-techmanual.dynamiclearningmaps.org/

Elementary and Secondary Education Act of 1965, Pub. L. No. 89–10, 79 Stat. 27 (1965). https://www.govinfo.gov/content/pkg/STATUTE-79/pdf/STATUTE-79-Pg27.pdf

Every Student Succeeds Act, 20 U.S.C § 6301 (2015). https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf

Matthews, G. J., & Harel, O. (2011). Data confidentiality: A review of methods for statistical disclosure limitation and methods for assessing privacy. *Statistics Surveys*, *5*, 1–29. https://doi.org/10.1214/11-SS074.

Nash, B., Clark, A. K., & Karvonen, M. (2016). *First contact: A census report on the characteristics of students eligible to take alternate assessments* (Technical Report No. 16-01). University of Kansas, Center for Educational Testing and Evaluation. https://dynamiclearningmaps.org/sites/default/files/documents/publication/First_Contact_Census_2016.pdf

Nash, B., & Thompson, W. J. (2017, April 26–30). *Evaluating an initialization tool for student placement into a map-based assessment* [Paper presentation]. National Council on Measurement in Education Annual Meeting, San Antonio, TX. https://dynamiclearningmaps.org/sites/default/files/documents/presentations/Initialization_Tool_for_Student_Placement_NCME_April_2017.pdf

*Recommended Waiver Authority Under Section 3511(d)(4) of Division A of the Coronavirus Aid, Relief, and Economic Security Act ("CARES ACT")*, 116th Cong. (2020) (Report to congress of U.S. Secretary of Education Betsy DeVos). https://www2.ed.gov/documents/coronavirus/cares-waiver-report.pdf

Rosenblum, I. (2021, February 22). [Letter to chief state school officers]. United States Department of Education, Office of Elementary and Secondary Education. https://www2.ed.gov/policy/elsec/guid/stateletters/dcl-assessments-and-acct-022221.pdf

Thompson, W. J., & Hoover, J. C. (2021). *Estimating changes in performance level distributions with a disrupted population* (Technical Report No. 21-01). University of Kansas, Accessible Teaching, Learning, and Assessment Systems. https://dynamiclearningmaps.org/sites/default/files/documents/publication/Propensity_Scores_Cross_Year_Comparisons.pdf