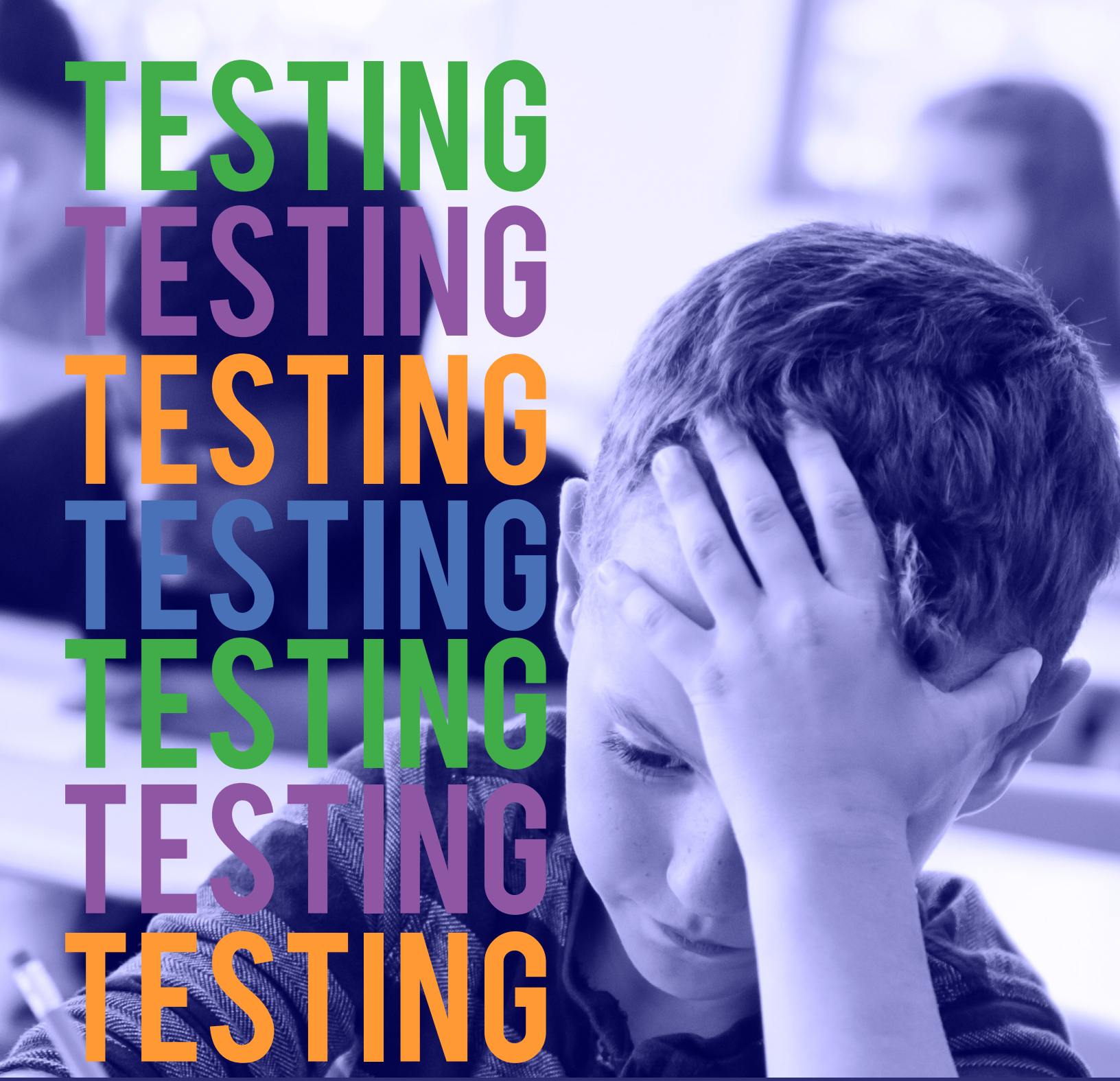


Understanding District & State Testing

Angela Bass & Rachel Canter

March 2018

MISSISSIPPI FIRST

ACKNOWLEDGMENTS & AUTHORS

This report was made possible by the generous support of the W. K. Kellogg Foundation. The views expressed in this paper are those of the authors alone.

We extend special thanks to teachers, testing coordinators, and administrators in our participating school districts.

Rachel Canter is the Executive Director. She is the author of all of Mississippi First reports, including the *State of Pre-K* and *Mississippi Voices*. Rachel founded Mississippi First in 2008.
rachel@mississippifirst.org

Angela Bass is the Deputy Director of Policy. She is the co-author of many of Mississippi First's reports, including the *State of Pre-K* and *Mississippi Voices.*
angela@mississippifirst.org

MacKenzie Stroh Hines is the Director of Communications. MacKenzie designs all Mississippi First materials, including this report, and leads Mississippi First's communication strategy.
mackenzie@mississippifirst.org

MISSISSIPPI FIRST

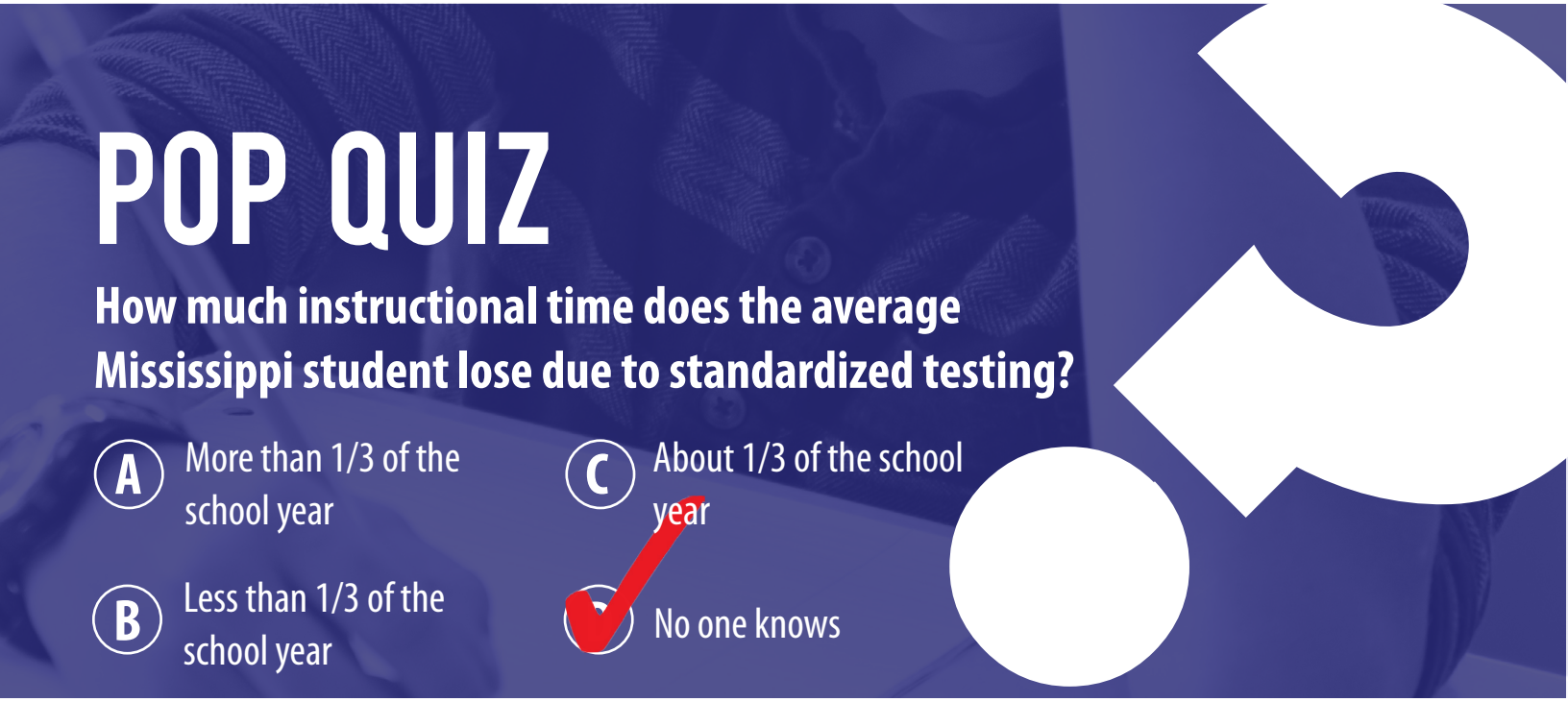
About Mississippi First

Mississippi First is a 501c3 public policy nonprofit specializing in improving public education in Mississippi. Mississippi First is a leading voice for state-funded pre-K, high-quality public charter schools, and rigorous state learning standards.

Report Art: iStockphoto
© Copyright 2018 Mississippi First
Mississippi First encourages the non-commercial use, reproduction, and distribution of our ideas, perspectives, and analyses. We require attribution for all use. Commercial use is not allowed.

125 S. Congress Street, Suite 1510
Jackson, MS 39201
601.398.9008 • www.mississippifirst.org

INTRODUCTION



Despite intense interest in the subject over the last few years, **Mississippi does not know how much instructional time the average student loses due to standardized testing.**

The lack of good data on "over-testing" has not stopped policymakers from trying to legislate against it. In January 2017, Representative Jeffrey Guice of Ocean Springs proposed a bill to limit the number of state testing days to three and the number of district benchmark assessment days to 20. *Mississippi Today* reported Representative Guice as saying, "We've all been told about standardized tests being administered between 40 and 60 days, [and there are] a lot of upset teachers and parents."[1] In 2015, the Mississippi Association of School Superintendents (MASS) pegged the number of testing days as between 38 and 45 of a 180-day school year in an article by the *Clarion-Ledger*.[2] For the same story, Lisa Karmacharya, the Executive Director of the Mississippi Association of School Administrators, described standardized testing as "a runaway train."

Tests may not be loved by either politicians or schoolchildren, but they are a necessary fact of both school and life. Good assessments help teachers determine what children know and can do and help parents ensure that their children are learning. Assessments also help policymakers evaluate the effectiveness of various programs and reforms and help the public understand the quality and equity of our education system. Finally, good assessments help students better understand what they have learned and build their ability to remember important information in the future.[3]

Despite the important educational role of assessments, they can become detrimental if used improperly. Every minute spent testing is one fewer minute for instruction. Additionally, testing can increase anxiety among students and teachers. Without knowledge about testing time in Mississippi districts, state leaders cannot ensure that the benefits derived from time spent testing outweigh the costs in lost instructional time and anxiety.

In this report, Mississippi First offers initial answers about the state of standardized testing in Mississippi public schools by examining testing practices in four diverse Mississippi school districts. We also provide broad recommendations for improving standardized testing and suggest action steps that school districts, the Mississippi Department of Education (MDE), and legislators should take to solve this problem.

[1] Royals, Kate. 2017. "House Panel Approves Shorter School Year, Less Testing." *Mississippi Today*, January 26. https://mississippitoday.org/2017/01/26/house-panel-approves-shorter-school-year-less-testing/.

[2] Mitchell, Jerry. 2015. "Standardized Tests Taking Toll on Mississippi Schools." *Clarion-Ledger*, March 1. http://www.clarionledger.com/story/news/2015/03/01/standardized-tests-taking-toll-mississippi-schools/24239399/.

[3] The Meadows Center for Preventing Educational Risk. 2017. *10 Key Policies and Practices for Assessment in Schools.* Research Summary, Austin, TX: University of Texas-Austin. Accessed November 15, 2017. https://www.meadowscenter.org/files/resources/10Key_Assessment_Web.pdf.

RESEARCH & SAMPLE

To learn more about standardized testing in Mississippi school districts, we conducted field research in four school districts from September to November 2015.

We sought to answer the following questions:

- **How much time do students in each grade in each district spend on district, state, and other significant standardized testing?**
- **What factors, if any, increase the amount of time students are exposed to standardized testing?**
- **What does the data suggest schools should or should not do to maximize the value of standardized testing?**
- **What are teachers' perceptions of standardized testing?**

To answer these questions, we adapted research tools from Achieve, a national education policy organization. These adapted tools include a Student Assessment Inventory completed by each district's administration and a focus group protocol that we used with elementary, middle, and high school teachers in each district. In addition, we developed a Technology Inventory with guidance from technology experts in Mississippi school districts to assess the technological capabilities in each district. More information on these tools can be found in Appendix A.

We recruited school districts to participate in this study based on their enrollment, accountability rating, poverty level, and per-pupil revenue. We also considered whether districts had 1-to-1 technology programs, which provide a laptop or computer for each enrolled child. In order to ensure we received candid responses, we granted anonymity to each district: all the names in this report are pseudonyms. We further provided an in-depth, district-specific report of the findings to each school district as an incentive for participating in this study. We gave each district a 21-day window to notify us of any errors to be corrected in those reports before we produced a final district report. We have only used data in this report from the final district reports.

Table 1 lists each district and gives approximate figures for each on the selection criteria.

TABLE 1. DISTRICT CHARACTERISTICS

District	Enrollment	2014-2015 Accountability Rating	FRL Percentage (2014-2015)[4]	Per-Pupil Revenue (2014-2015)	1-to-1 Technology?
Overton	~1,500	D	>95%	~$11,000	No
Hillside	~2,700	A	~60%	~$8,700	Yes
Mannequin	~5,800	A	~40%	~$8,500	No
Sunset	~5,600	D	>95%	~$8,900	Partial

[4] "FRL percentage" stands for "free and reduced-price lunch percentage," a measure of poverty using rates of students qualifying for free or reduced-price school lunch. Qualifying students come from homes at or below 185% of federal poverty guidelines.

CALCULATING TESTING TIME

This research sought to quantify the number and types of tests that ***all students*** in a particular grade take.

In calculating testing times, we excluded tests administered only to a subset of students, such as placement testing for children with individualized education plans or for English language learners, or tests students choose to take as part of a course of study, such as AP tests. We also excluded tests administered to a random statistical sample for research purposes or tests administered on a periodic basis, such as the National Assessment for Educational Progress (NAEP), which is only administered to a sample of students every few years. Other tests that are not included in this study include teacher-created tests not given schoolwide and nine-weeks tests not given districtwide. We discuss these other tests in the case studies of each district in Appendix B to provide further context to the universe of testing in each district.

After these exclusions, tests fell into two main categories—state-mandated and district-mandated tests.

State-mandated tests include all tests required by state law or policy, such as those required for accountability purposes. These are the *Mississippi K-3 Assessment Support System (MKAS^2)* that kindergarten and third grade students take;[5] the *literacy screeners* that kindergarten through third graders take; the *state exams in English/Language Arts (ELA)* and *math* that third through eighth graders take; the *state science exams* that fifth and eighth graders take; the *subject area tests* in *Algebra I, Biology I, English II,* and *US History* that high school students take; and the *ACT* that all high school juniors take.

To calculate state testing time, we use official times for each test provided by MDE. In 2014-2015, all state tests were administered via computer except for the ACT. The 2014-2015 school year was the only year that Mississippi took the Partnership for Assessment of Readiness for College and Careers (PARCC) test. All state testing times in this report include PARCC times unless otherwise noted. We discuss how the 2017-2018 Mississippi Academic Assessment Program (MAAP) test changes our findings, as relevant, in the sections below. Please see Appendix C for a master table of state assessment times.

District-mandated tests vary by district, but in our sample, they mostly consist of *progress-monitoring exams*, which provide information to teachers and leaders about students' progress throughout the year and their expected performance on state-mandated tests. School district testing policies are set by district administrators and local school boards, which have the flexibility to require or limit assessments as they see fit.

To calculate district testing time, we used our Student Assessment Inventory to determine which tests school districts administer to all students in each grade. Most, but not all, district-mandated assessments in our study were computerized. When schools used vendor-created products, such as the i-Ready assessments, we use official vendor testing times to ensure consistency across districts. When schools reported using district-created assessments, we rely on district-reported test times. Appendix C also contains a master table of all district assessment times from this study.

Please note: The testing times in this study only reflect time for ***actual test completion*** as defined by the amount of time that students are answering test questions. Activities such as logging in for computer-based tests or passing out test booklets are not factored into our numbers. For untimed tests, we have either used a district- or state-reported range or average test time. Remember that this data represents the ***minimum amount of instructional time lost for the purposes of standardized assessment.*** Actual time lost may be much greater, depending on the efficiency of total test administration. We address these issues in General Finding 3.

[5] Recently, the State Board of Education voted to discontinue the use of the MKAS^2 for third graders beginning in 2018-2019. MKAS^2 will continue to be the state's kindergarten-entry assessment.

GENERAL FINDINGS

In this section, we present findings true of all the districts we studied. These findings are grouped by the following categories—***tests and testing time***, ***district testing choices***, ***teacher perspectives***, and ***student and parent perspectives.***

TESTS AND TESTING TIME

FINDING 1: In 2014-2015, students spent an average of 7 hours, 53 minutes—less than 1% of a 180-day school year—taking state tests. We also analyzed time spent taking state tests by grade span (see Table 2) and found that no grade span spent more than 1.1% of the school year taking state tests. Eighth grade students spent the most time completing state ELA, math, and science tests, at 13 hours and 50 minutes. Even for these students, state testing only amounted to just over two six-hour school days, or a little more than 1% of the school year. With such a small percentage of time devoted to taking state tests, the frustrations we heard from teachers about state testing time may seem strange, but our research found that these frustrations reflected a real increase in state testing time in 2014-2015, the only year that Mississippi students took the PARCC test for ELA and math.

The PARCC test added a significant amount of time to annual state testing. PARCC required two administrations—a performance-based assessment occurring after 75% of the school year and a summative assessment occurring at the end of the year—per tested grade in ELA and math. In ELA, the performance-based assessment (PBA) consisted of two or three timed parts, ranging in duration from 60-90 minutes. In math, the PBA consisted of two timed parts, ranging in duration from 70-90 minutes. Each end-of-year test in ELA and math consisted of two timed parts, ranging in duration from 60-90 minutes.

TABLE 2. Average Time Spent on State Tests by Grade Span

Grade	Hours, Minutes	Percentage of School Year
K-2	1 hour, 30 minutes	<1%
3-8	11 hours, 41 minutes	1.1%
9-11[6]	7 hours	<1%

The addition of the PBA meant that PARCC required one more test administration for the high school ELA and math exams and two more test administrations in grades 3-8 than any previous state test or the current state MAAP exam. Districts that had to rotate students through computer labs due to a lack of technology had to endure this process twice—once for the PBA and once for the end-of-year exams. Even districts without technology issues faced two state testing

[6] When students take the four state-mandated subject area tests in high school is dependent on when they take the course tied to the exam. For example, a small percentage of students take the Algebra I exam in eighth grade while some might take it in tenth, eleventh, or twelfth grades. Generally, however, Algebra I is viewed as a ninth-grade course with a corresponding exam. We have assumed in our calculations that most students take the subject area tests on the following schedule: ninth grade—Algebra I and Biology I, tenth grade—English II, and eleventh grade—US History and the ACT.

TESTS AND TESTING TIME, CONTINUED

windows and all of the attendant disruption. For teachers, the PARCC testing experience exacerbated discontent about a perceived lack of communication around what students and teachers could expect from PARCC. To add insult to injury, teachers did not receive PARCC scores until almost a year after PARCC was administered.

2014-2015 was the only year that Mississippi took the PARCC test. In 2015-2016, Mississippi children took the Questar-designed Mississippi Assessment Program (MAP) test for ELA and math. MAP was a shorter test because it required an end-of-year administration only, even though each test was actually longer than the PARCC end-of-year tests. MAP's end-of-year tests consisted of two parts per subject: an online test and a performance task. In ELA, the online test ranged from 113-138 minutes depending on grade level, while the performance task[7] ranged from 90-100 minutes. In math, the online test ranged from 127-157 minutes, while the performance task ranged from 70-90 minutes. MAP reduced state testing time in 2015-2016 by an average of 2.68 hours per MAP-tested grade and, consequently, reduced total state testing time.

In 2017-2018, MDE updated the testing times and renamed the test "MAAP," a name that now encompasses all state tests, not only ELA and math. Testing time for math remained the same, but testing time for ELA increased. Now, all grade levels take a 150-minute performance task[8] in addition to the 113-138-minute online test. This change increases total testing time by 60 minutes in grades 3-8 and 50 minutes in high school. As a result, MAAP reduces total state testing time compared to 2014-2015 an average of 1.83 hours per MAAP-tested grade (see Table 3). Including all 2017-2018 state tests, students take the greatest number of state-mandated tests in grade 3, with four state tests and a total of six administrations. Eighth grade students still spend the greatest amount of time on state testing (11.92 hours). Nonetheless, no grade spends more than approximately 1% of the year completing state tests.

TABLE 3. State Test Completion Time in Hours per Grade, 2014-2015 v. 2017-2018[9]

Year	K	1	2	3	4	5	6	7	8	9	10	11	12
2014-2015	2.7	1	1	11.58	10	13	10.83	10.83	13.83	7.33-9.33	5.75	4.92-6.92	0
2017-2018	2.7	1	1	9.5	7.67	10.67	8.92	8.92	11.92	6.12-8.12	4.8	4.92-6.92	0
Change	0	0	0	-2.08	-2.33	-2.33	-1.91	-1.91	-1.91	-1.21	-.95	0	0

[7] The 2015-2016 ELA performance task was a paper-and-pencil assessment. All other MAP tests were administered online.
[8] The 2017-2018 ELA performance task is given online. All MAAP tests are now administered online.
[9] This table includes testing time for all state tests, not just PARCC or MAAP ELA and math.

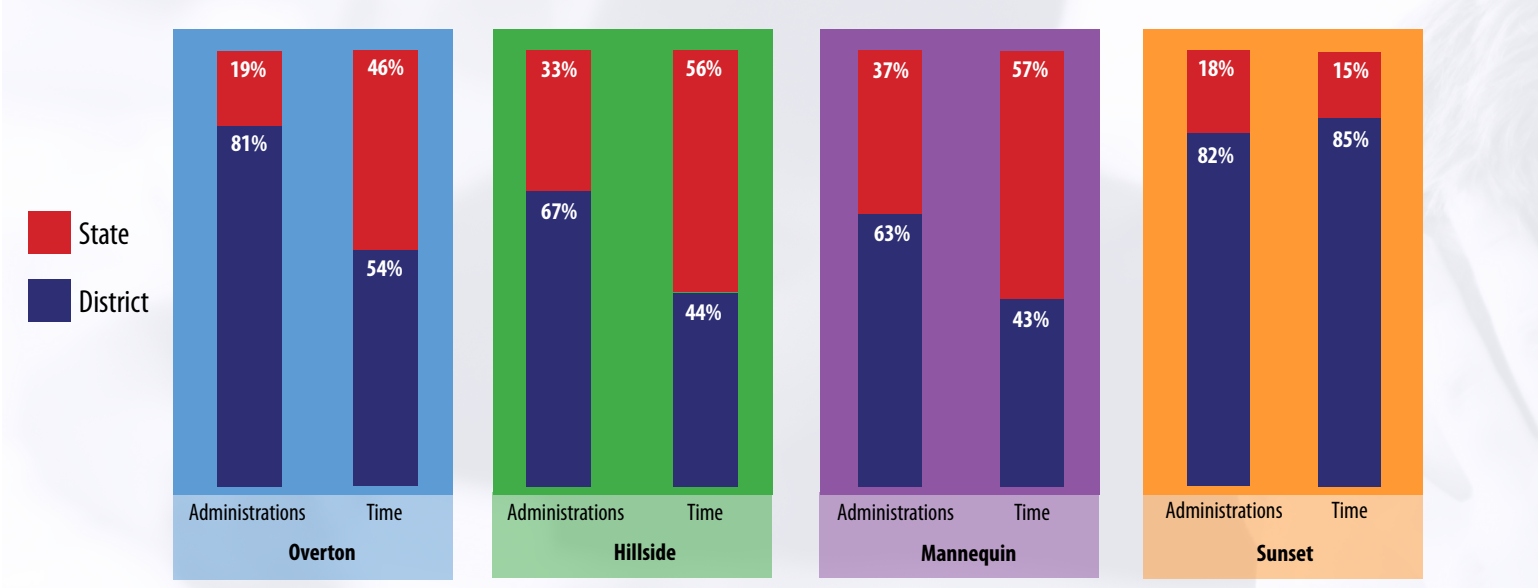
TESTS AND TESTING TIME, CONTINUED

FINDING 2: Students took more district tests than state tests in every district we studied, but they sometimes spent less time on district testing than on state testing. Every district we studied administered standardized tests not required by the state. At least 63% of all standardized tests taken by students were district tests. Although district tests were administered more frequently, this did not necessarily translate into more district testing time, at least in terms of test completion time. Two districts—Hillside and Mannequin—spent less time overall on district testing (approximately 44% and 43%, respectively). The other two districts—Overton and Sunset—spent more time overall on district testing (approximately 54% and 85%, respectively).

For each district, many factors influenced whether district or state testing took more time in terms of hours students spent completing tests. On average, a state test—whether PARCC or MAAP—is longer than a district assessment in each of the districts we studied (even with Sunset's long nine-weeks tests skewing the average). The districts with greater district testing times, therefore, also tested more frequently to result in higher overall testing times. We describe this finding further in Comparative Finding 2.

Figure 1 shows the percentage of district test administrations compared with state test administrations as well as the percentage of total time attributed to district tests versus state tests per district. For most of these comparisons, presented side-by-side below, state testing time is more evenly split with district testing time than state administrations are with district administrations.

FIGURE 1. Standardized Test Administrations and Time, District v. State



TESTS AND TESTING TIME, CONTINUED

Finding 3: Test completion hours do not reflect all the time schools devote to standardized testing.
For each assessment, schools must weigh the value of the feedback to students and teachers against the cost of lost instructional time. Extensive research shows that thoughtful testing practices help students deepen their understanding and retain what they have learned.[10] Shorter, frequent assessments, when used in the context of a well-designed curriculum, can even be more beneficial to students than a lone summative assessment.

Lost instructional time comes not only from the time it takes to complete a test but also from testing-related disruptions to the normal school day. If severe, these disruptions could cause three one-hour tests to actually take more time than one three-hour test. This is why the large number of district test administrations is concerning, even though state test completion time may be longer each year in some districts. For frequent assessments to increase learning, schools must minimize testing-related disruptions and ensure students and teachers utilize the feedback that the assessments provide.

This report explicitly did not try to quantify all of the time that students, teachers, and schools lose from factors other than test completion. These testing-related disruptions include test prep, test administration protocols, day-of-test inefficiency, schedule changes, and loss of productivity after the test. When these disruptions add up, they can prevent testing days from also being learning days, even if only a small part of the day is spent completing a test. Data from our assessment inventory and focus groups is illuminating in understanding the impact of these other factors.

Before students can begin taking tests, schools must send children to testing rooms, pass out testing materials, or log students in to computers. Depending on the type of test—and how strictly each of these activities must be controlled—these activities may require more or less time. For state testing, schools must follow very specific test security protocols, which require teachers and students to behave in formal, ritualized ways to ensure that every student statewide has the same testing experience. Typically, these protocols include directions related to handling and distributing test materials, taking up cellphones, checking identification (for the ACT), or logging in to computers, as well as what to say before students begin the test. State testing protocols likely add more time pre-test than less formal activities required before a student takes a progress-monitoring test.

Testing-related disruptions include **test prep**, **test administration protocols**, **day-of-test inefficiency**, **schedule changes**, and **loss of productivity after the test**.

Inefficiency at the district and school levels can also impact the amount of time that schools experience testing-related disruption. Teachers in Overton described preventable problems with state test administration leading to longer testing days. Similarly, teachers and district personnel in Sunset listed "building preparation" for the district nine-weeks tests as a cause of longer testing days. Many districts also noted that "unplanned technical issues" could impact testing time for computerized tests. These issues may be problems with the computers used for testing, problems with the testing platform itself, or even problems with students forgetting passwords needed to login. Availability of technology also plays a role in test administration efficiency, as described more in Comparative Finding 4.

Particularly long tests, like state tests or district nine-weeks tests, require schools to change their schedules, leading to a loss of instructional time for non-tested classes. Schools may accommodate this extra time either by shortening each class for the remainder of the day or not sending students to particular classes at all. Teachers in elementary schools commonly reported that students skipped their special classes (e.g., P.E. or music) during testing periods, either to focus remaining time on core classes or to use those teachers as proctors. State testing requires special schedules not only because students need longer than a typical class period to finish tests but also because every tested student must finish before the school

[10] The Meadows Center for Preventing Educational Risk. 2017. *10 Key Policies and Practices for Assessment in Schools*. Research Summary, Austin, TX: University of Texas-Austin. Accessed November 15, 2017. https://www.meadowscenter.org/files/resources/10Key_Assessment_Web.pdf.

TESTS AND TESTING TIME, CONTINUED

can resume operations. Sunset, which had the highest testing time among the districts we studied, also follows this practice for its nine-weeks testing, which teachers described as the district going on "lock-down." Sunset teachers reported sitting idle in classrooms for much of the district testing block simply waiting on other students in the building to finish. Because it must accommodate four three-hour tests every nine weeks, Sunset loses an entire week to testing each quarter. Time not spent testing during that week is spent reviewing, as teachers are directed not to teach new material for the week.

Progress-monitoring assessments are less likely to cause disruptive schedule changes as they are usually designed to take no more than one 50-minute class period, although some vendors suggest that schools split hour-long tests between two 50-minute class periods. The computerized nature of these tests allow students to pause and resume the test with relative ease, and schools worry less about test security if students have to finish on a subsequent day. At least one district—Mannequin—tries to prevent instructional time loss in ELA and math by requiring progress-monitoring tests to occur in social studies or science classes. This protects time in ELA and math but reduces it in another core subject, a frustration for those teachers.

We also found that schools usually do not administer standardized tests back-to-back in the same day. Tests in different subjects are spread across several days, and multi-part exams in the same subject are often taken on subsequent days. This is nearly always true of state testing and frequently true of district nine-weeks testing. For instance, instead of administering the two-part MAAP English II assessment that calls for four hours and 48 minutes on one day, districts may administer part one of the assessment on one day and part two on the next day, taking two hours and eighteen minutes the first day and two and a half hours the second day. By spreading tests over multiple days, schools must repeat pre-test activities and wait for all students to finish testing each time, if required as part of a protocol.

Adding up all the days that **any child in any grade somewhere in the district** is taking either a district or state test could easily reach the 38-45 days cited by MASS, but this is a **very different number** than the number of days that **an individual student** is taking a district- or state-mandated test.

Finally, teachers noted that all tests, especially long tests, impact the productivity of the classroom immediately after testing finishes. This problem is most acute for state testing. Teachers across our sample revealed that once state testing is over for the day or the week there is very little expectation that either tested or non-tested students participate in instructional activities, even if plenty of the school day remains. Teachers in the Mannequin district explained this practice, stating: "Teachers try to teach [after a test], but we have found that students need a mental break." A teacher in the Hillside District also noted, "By the time children reach the final test on a multi-day test, they are fatigued." This finding is also true of long district tests.

All of the aforementioned factors explain why anecdotal reports have higher estimates of testing time than what this study reports. Many teachers in our focus groups included the number of days that standardized tests are administered in a school instead of the actual time that students spent completing tests when talking about testing time. They also do not differentiate between state or district testing in these comments, which may give laypeople the mistaken impression that they are referring to state testing only. Adding up all the days that any child in any grade somewhere in the district is taking either a district or state test could easily reach the 38-45 days cited by MASS, but this is a very different number than the number of days that an individual student is taking a district- or state-mandated test. Furthermore, the number of days on which a district administers any test can be wildly divergent across districts. This is likely due to the many contextual factors above. For example, each of our districts listed a different number of days required to test all third graders on the state-required MKAS^2. Hillside reported 1 day, Sunset reported 5 days, Mannequin reported 5-10 days, and Overton reported 9-14 days. For an individual third grader, the MKAS^2 is only 50 minutes long.

TESTS AND TESTING TIME, CONTINUED

	District	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May
K	Sunset	3	0	0	0	2	0	0	0	0	3
	Overton	0	5	2	2	2	4	2	4	2	3
	Hillside	4	0	2	2	1	2	0	2	0	4
	Mannequin	3	0	2	0	2	2	0	2	0	5
1	Sunset	2	0	0	0	2	0	0	0	0	2
	Overton	0	4	2	2	2	4	2	4	2	2
	Hillside	2	0	2	3	0	2	0	2	0	2
	Mannequin	2	0	2	0	2	2	0	2	0	4
2	Sunset	2	0	4	0	6	0	0	4	0	6
	Overton	0	4	2	2	2	4	2	4	2	2
	Hillside	3	0	3	3	0	3	0	3	0	3
	Mannequin	2	0	2	0	2	2	0	2	0	4
3	Sunset	2	0	4	0	6	0	0	6	0	9
	Overton	0	4	2	2	2	4	2	6	2	5
	Hillside	3	0	3	3	0	3	0	5	0	5
	Mannequin	1	0	2	0	2	1	0	4	0	6
4	Sunset	2	0	4	0	5	0	0	6	0	8
	Overton	0	4	2	2	2	4	2	6	2	4
	Hillside	3	0	1	1	0	1	0	3	0	5
	Mannequin	2	0	2	0	2	2	0	4	0	6
5	Sunset	2	0	4	0	6	0	0	6	0	9
	Overton	0	4	2	2	3	5	2	7	2	5
	Hillside	2	0	0	0	2	0	0	2	0	5
	Mannequin	2	0	2	0	2	2	0	4	0	7
6	Sunset	1	0	4	0	5	0	0	6	0	7
	Overton	0	4	2	2	2	4	2	6	2	4
	Hillside	1	0	0	0	2	0	0	2	0	4
	Mannequin	2	0	2	0	2	2	0	4	0	6
7	Sunset	1	0	4	0	5	0	0	6	0	7
	Overton	0	4	2	2	2	4	2	6	2	4
	Hillside	1	0	0	0	2	0	0	2	0	3
	Mannequin	2	0	2	0	2	2	0	4	0	6
8	Sunset	1	0	4	0	5	0	0	6	0	8
	Overton	0	4	2	2	2	5	2	7	2	5
	Hillside	1	0	0	0	2	0	0	2	0	7
	Mannequin	2	0	2	0	2	2	0	4	0	7
	Average	1.50	1.03	2.08	0.83	2.44	1.83	0.50	3.97	0.50	5.06

FINDING 4: Certain times of the year are more affected by testing. According to our data, students in K-8 experience the heaviest testing in May due to multiple state tests as well as end-of-year district testing (see Table 4). October, December, and March are also high testing times, due to nine-weeks testing and progress-monitoring exams. In 2014-2015, March had unusually intense testing due to the PARCC performance-based assessments. High school students who are taking state-tested courses may be testing frequently at the end of either semester, if on a 4x4 block, or in May, if on a period schedule.

TABLE 4.

Number of K-8 Test Administrations (State and District) in Each School District by Month

GENERAL FINDINGS, CONTINUED

DISTRICT TESTING CHOICES

FINDING 5: Districts rely on vendor-created standardized testing products. All districts in our sample used testing products created by at least one vendor and as many as six. Every district, for example, used the STAR Reading and STAR Math products. The STAR products are owned by Renaissance Learning, the national testing vendor who designed the MKAS^2 exams. In 2014-2015, MDE used the MKAS^2 as both the kindergarten-entry assessment and the third grade reading promotion exam, also called the third grade reading "gate." Since 2016-2017, MDE has used the reading and language portions of the state ELA exam as the third grade gate instead of the MKAS^2, although the MKAS^2 continued to be administered and was used as the re-test. MDE also paid for and required each district to use the STAR Reading test as their K-3 literacy screeners.[11] This requirement drove many districts to independently adopt and pay for STAR Math. In addition to STAR, three of four districts used products by at least one more vendor. For two of these districts, this second vendor product was i-Ready.

Most vendor exams used by our sample are skills-based progress-monitoring exams that can be paired, in theory, with any curriculum using the state standards. These external, non-curricular testing products form the basis of most of our sample schools' data-driven decision-making. Only Hillside used vendor-created, curriculum-based exams for some grades.

Table 5 shows the vendor-created testing products that each district used (we did not include standardized tests created by district personnel in Table 5).

TABLE 5. Vendor-Created Testing Products Used by Districts

		ELA	Math	Science	Social Studies	Vendors/Products
Overton	Elementary	STAR Reading i-Ready ELA	STAR Math i-Ready Math	Case21 Science		3 vendors; 5 products
	Middle	STAR Reading i-Ready ELA	STAR Math i-Ready Math	Case21 Science		
	High	No district-mandated tests in high school				
Hillside	Elementary	STAR Reading DIBELS Scholastic Reading Index NWEA MAP ELA	STAR Math NWEA MAP Math			6 vendors; 18 products
	Middle	NWEA MAP ELA Curriculum ELA*	NWEA MAP Math Orleans-Hanna Curriculum Math I*	Curriculum Science*		
	High	Curriculum ELA I* Curriculum ELA II*	Curriculum Math II*	Curriculum Biology* Curriculum Chemistry*	Curriculum World History* Curriculum American History*	
Mannequin	Elementary	STAR Reading i-Ready ELA	STAR Math i-Ready Math			2 vendors; 4 products
	Middle	i-Ready ELA	i-Ready Math			
	High	No vendor-created tests in high school				
Sunset	Elementary	STAR Reading	STAR Math			1 vendor; 2 products
	Middle	STAR Reading				
	High	STAR Reading				

*To protect the anonymity of Hillside, we have not named the specific testing product used.

[11] When MDE recently "de-scoped" its contract with Renaissance to eliminate the MKAS^2 for third graders in 2018-2019, they also decided to stop paying for STAR Reading for all districts. Beginning in 2018-2019, districts will select a literacy screener from a list, as allowed by law.

DISTRICT TESTING CHOICES, CONTINUED

Even when districts used the same testing products, they used them very differently. Every district in our sample used the STAR math and literacy exams, as discussed above. However, each district used STAR very differently, as shown in Table 6 below.

Hillside and Mannequin used STAR in K-3 but preferred other testing products in grades 4-12. Overton and Sunset chose to continue STAR testing through middle school or, in Sunset's case, through graduation. Overton not only used STAR most frequently but was also the only district to use a second progress monitoring tool, i-Ready, in all of the same grades, although not as frequently (only three times per year). Overton's use of STAR every four weeks raised a red flag for us because testing products have guidelines about how frequently to use them to retain their validity and reliability. Renaissance Learning was careful to avoid telling us that an every-four-weeks' administration compromised the STAR test's validity or reliability, but they stated that three to six times per year is optimal.[12]

FINDING 6: District testing choices are strongly linked to state testing. Though districts have control over the number and types of district-mandated assessments they administer, data from our investigation suggest that district testing decisions are driven by trends in state testing.

All districts in our sample tested in both English and math at some or all grade levels. Both Mannequin and Hillside tested ELA and math K-10. Overton tested ELA and math K-8, while Sunset tested in grades K-12. Two districts—Mannequin and Overton—administered more math tests while the other two administered more ELA tests. In the districts administering more math tests, this was due to the K-3 literacy screeners counting as state tests rather than district tests, not a special focus on math.

Three of the four districts also administered standardized science and/or social studies tests at some or all grade levels. Sunset mandated testing in science and social studies in grades 2 through 12. Hillside only mandated science testing in grades 8-10 and social studies testing in grades 9-10. Overton mandated district science testing in grades 5 and 8. Mannequin did not mandate any district science or social studies testing.

TABLE 6. STAR Reading Administration Frequency

School District	Grades	Frequency
Overton	K-8	Every 4 weeks (9 times per year)
Hillside	K-3	Every 6 weeks (6 times per year)
Mannequin	K-3	Every 12 weeks (3 times per year)
Sunset	K-12	Every 12 weeks (3 times per year)

Figure 2 (page 12) shows the number of district-mandated standardized test administrations by subject area in each of our districts.

[12] Renaissance Learning, telephone conversation with A. Bass, September 1, 2017.

DISTRICT TESTING CHOICES, CONTINUED

As state testing in a grade increases, so does district testing in those grades. On average, students in the districts in our sample are exposed to the greatest number of district tests from grades 2-8. This tracks very closely to state exams, which are highest in grades 3-8. This average does hide some variations.

Hillside's administrations peak early from grades 2-3, while Overton's and Mannequin's administrations peak between fourth and eighth grade. Sunset's administrations are fairly consistent in grades 2-12, though they do peak in grades 4 and 5.

Figure 3 shows the number of district-mandated standardized tests by grade in each of the school districts in our sample as well as the four-district average.

FIGURE 2. District-Mandated Test Administrations per Subject by District

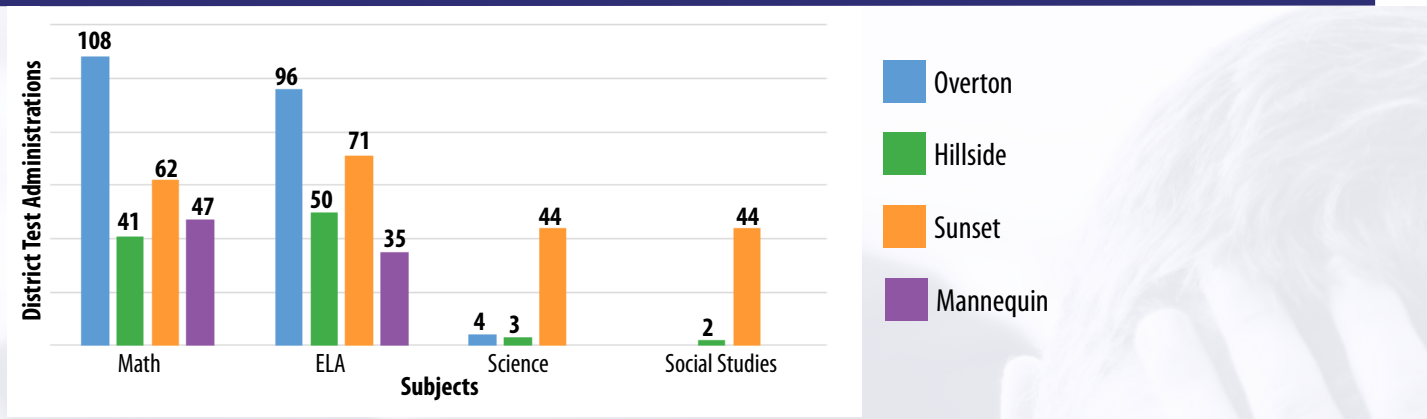
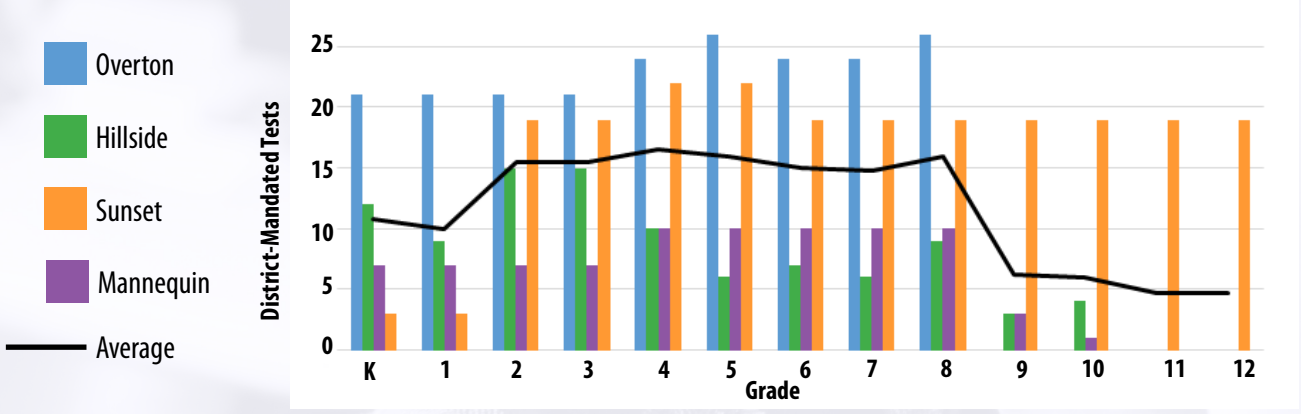


FIGURE 3. Number of District-Mandated Tests per Grade by District



GENERAL FINDINGS, CONTINUED

TEACHER PERSPECTIVES

FINDING 7: Teachers evaluated tests according to four criteria: relevance, timeliness, usability of data, and affiliated support for data use. Each of these criterion contributes to a test's usefulness to instruction, the most important purpose of testing for teachers. Though the number of test administrations and testing time varied among the districts in our sample, these criteria held true for all of the teachers participating in our focus groups.

Relevance—Teachers across districts found the relevance of a test to their instructional decision-making the most important factor influencing their opinions of the test. They judged relevance by the perceived alignment between what the test measures and what students have been taught or should have been taught. When teachers perceived tests to lack relevance, they were especially frustrated at being mandated to administer them.

- Teachers across the districts found *their own tests* to be the most relevant as they could write them based on exactly what they taught. Experienced teachers at some schools also felt that their own exams were more rigorous than either state or district tests and therefore more valid indicators of student knowledge. Teachers in Mannequin were the only ones to describe a structured process for writing weekly assessments as part of a grade-level team of teachers. These assessments, though they were common to all students in a grade, got equally high relevance marks from teachers.
- *District-created assessments* suffered on the relevance criterion in both districts that required district benchmark tests. Mannequin and Sunset teachers described a misalignment between pacing guides and tested objectives, leading to considerable frustration for both teachers and students when test questions covered material that had not been taught. In both of these districts, these tests were created by district personnel rather than teachers.

Teachers evaluate tests by **relevance**, **timeliness**, **usability of data**, and **affiliated support for data use**.

- Teachers across our sample had mixed reviews of the relevance of *tests used for progress monitoring*, such as STAR and i-Ready. Elementary teachers were generally positive about progress-monitoring tests. Middle school teachers, however, had more mixed reviews. Some middle school teachers found these tests to be helpful while others were not sure whether they were helpful enough to justify the time spent testing. In one district, out-of-subject teachers were asked to administer the tests (e.g., science teachers asked to administer reading tests), which was a sore spot with them as the results were not relevant to their subject. In another district, middle school teachers did not ever receive the data for the tests they administered in their subject area. Some middle school teachers and all high school teachers reported widespread problems with students not taking progress-monitoring tests seriously, making the results not very relevant to instruction, even if they deemed the material tested relevant.
- *Curriculum-based exams*, such as those used in Hillside, received extremely high relevance reviews. Teachers could see the direct results of their teaching in how students performed on summative curriculum tests created by the curriculum vendor. For Hillside teachers, these end-of-the-year tests were what state tests should aspire to be. Teachers in Hillside, however, worried whether the curriculum exams were aligned to state tests. They feared that even if students did very well on the summative curriculum exams, their knowledge might not translate to state tests, which could be testing slightly different skills or content.
- Finally, teachers across districts found *state tests* to be relevant in the sense that they reflected state standards, but they did not always find them rigorous. Teachers noticed the increase in rigor with PARCC, but high school teachers in Mannequin felt that the rigor of the U.S. History and Biology I exams was still lacking.

TEACHER PERSPECTIVES, CONTINUED

Timeliness—Teachers strongly preferred administering tests that produced timely results—the quicker, the better. Computerized tests that report feedback instantly were most preferred. Generally, teachers wanted to receive data shortly after a test was administered and use it to direct their practice right away. This was a strength of vendor-created progress-monitoring exams, which are taken via computer and report results instantly.

Teachers' views of a test's timeliness were also influenced by what type of test it was. Across our sample, teachers preferred the results of tests used for progress monitoring immediately (24-72 hours) or their value declined sharply. The results of summative exams, such as state tests, could take a few weeks. The important point is that the state results arrive before teachers are too far along in planning for—or teaching!—the next school year. From the instructional perspective, state test results are only useful if they can direct a teacher's practice for the following year's students, and the long lag between test administration and receipt of results means there is usually little focus on making sure teachers can use the data in this way. Teachers across districts thought of state tests as "autopsies" after the students have moved on. As one middle school teacher stated, "State testing is helpful for the state, so that they can grade schools, but it is not helpful for me." This statement summarized the feelings of teachers across our sample who saw the value of state tests only as an accountability measure or as a measure for parents and students to understand their end-of-year progress. Many teachers complained particularly about the timeliness of the results from the state PARCC assessments. As of our last focus groups in November 2015, teachers had not yet received PARCC scores. On the flip side, while MKAS[2] results had a relatively short turnaround for a state assessment (two weeks), teachers in Mannequin felt that this was still not quick enough because the re-test timeline was so abbreviated.

> "State testing is **helpful for the state**, so that they can grade schools, but it is **not helpful for me**."

Usability of data—Even if tests were highly relevant and data was timely, if the data was not in an immediately usable format, teachers across the districts in our sample found the process of reformatting the data to be burdensome and, therefore, the test less valuable. These teachers wanted to be able to "turnkey" assessment results, which led to a preference for computerized exams that create sophisticated reports by objective and by student. Assessments in use in each of the districts, such as STAR and i-Ready, have this capability and can even automatically assign students to remediation groups.

Affiliated support for data use—Teachers had more positive views of any particular test if there was a formal structure in place for data analysis and discussion to help teachers improve their practice. Teachers in Mannequin, for example, highly valued the district benchmark data meetings where they could compare scores with other teachers and learn from fellow teachers' expertise, even though Mannequin teachers did not favor the district benchmark tests. In Hillside, teachers liked the ability to meet with administrators and coaches one-on-one to discuss data from their end-of-year curriculum exams, which return results in August. The strength of this data support in Hillside seemed to balance out the fact that Hillside's curriculum tests are less timely than a product like STAR or NWEA MAP.

GENERAL FINDINGS, CONTINUED

STUDENT AND PARENT PERSPECTIVES

FINDING 8: Teachers report widespread confusion among students and parents as to whether tests are district or state mandated and why they are important. In focus groups, teachers across each of the districts in our sample described an environment in which standardized testing had become so frequent that parents and students could not differentiate the tests. Teachers reported that many parents did not know the difference between state and district standardized assessments or why some tests may be more significant than others. Additionally, teachers reported that students also did not understand the purpose behind tests, especially assessments that had little to do with their grades. Reported reactions from students and parents tended to vacillate between two extremes—severe angst and apathy. This finding was true regardless of the performance of the district. In Hillside, one of the high-performing districts, one teacher related a story about a student scheduling a dental appointment during an important test because the parent thought it was "just another school day."

In Review: GENERAL FINDINGS

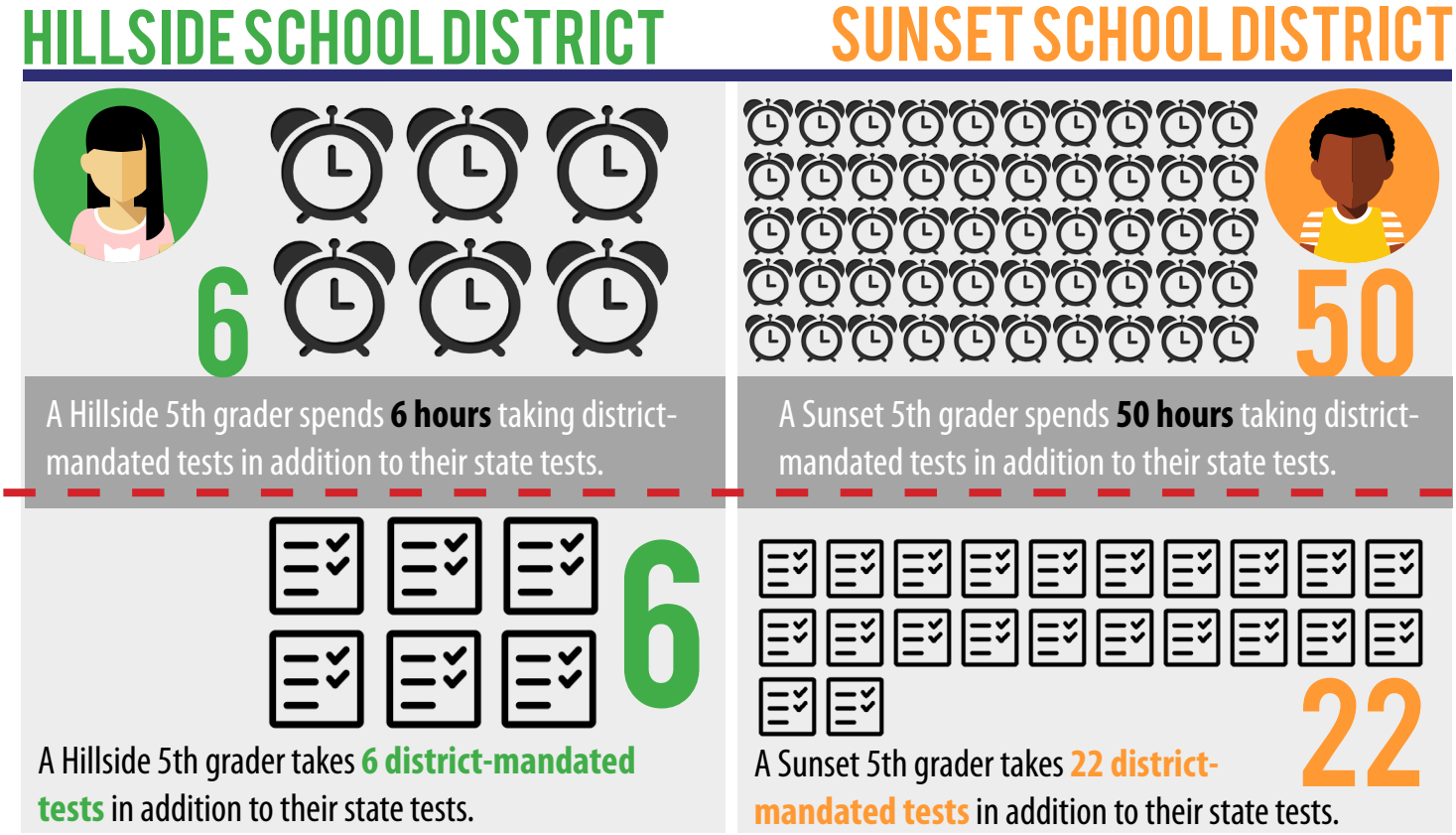
1. In 2014-2015, students spent an average of 7 hours, 53 minutes—**less than 1% of a 180-day school year**—taking **state tests**.
2. Students took more **district tests than state tests** in every district we studied, **but** they sometimes spent less time on district testing than on state testing.
3. Test completion hours **do not reflect all the time** schools devote to standardized testing.
4. **Certain times of the year** are more affected by testing.
5. Districts rely on **vendor-created standardized testing products.**
6. District testing choices are **strongly linked to state testing.**
7. Teachers evaluated tests according to four criteria: **relevance**, **timeliness**, **usability of data**, and **affiliated support for data use**.
8. Teachers report **widespread confusion** among students and parents as to whether tests are district or state mandated and why they are important.

COMPARATIVE FINDINGS

In this section, we compare differences in standardized testing across the districts in our sample.[13]

COMPARATIVE FINDING 1: Districts have wide autonomy in how they use testing, resulting in very different student experiences across districts. Each student in the state takes the same number of state tests, but because each district controls the amount of district testing, students in some districts are exposed to more standardized testing than others. We found that the amount of district-mandated testing ranged from an average of 10 tests per grade to 27 tests per grade, depending on the district in our sample. For instance, a fifth grade student in the Hillside District spends a total of six hours taking six district-mandated standardized assessments in addition to state tests. In the Sunset District, a fifth grade student spends 50 hours taking 22 district-mandated standardized exams in addition to state-mandated exams. This finding has great implications for the number of lost instructional hours in each district.

FIGURE 4. COMPARISON OF TESTING TIMES & NUMBER OF TESTS



[13] We highlight similarities and differences based on district characteristics, but readers should note that the findings from districts in a category are not necessarily representative of all districts in that particular category statewide due to our small sample size. For instance, our findings related to low-performing school districts cannot be generalized to all low-performing school districts without further research. Instead, readers should view this information as suggestive of avenues for further inquiry.

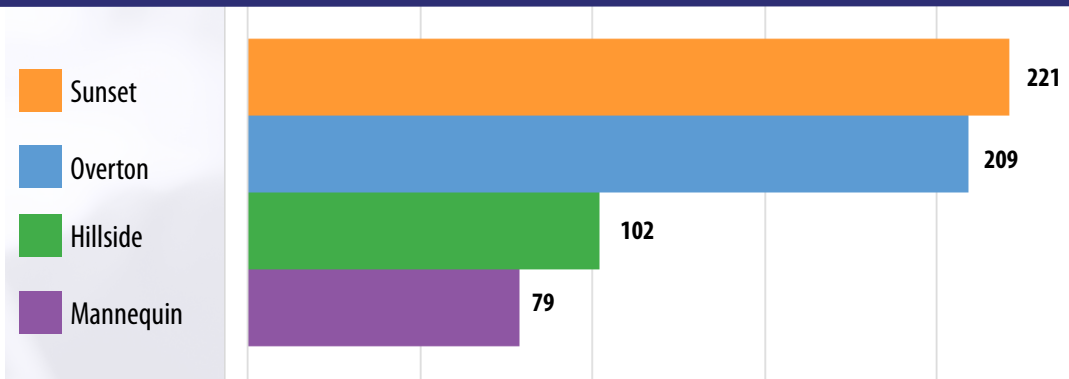
COMPARATIVE FINDINGS, CONTINUED

COMPARATIVE FINDING 2: Low-performing districts in our sample administered more tests and spent more time testing than high-performing districts. Overton and Sunset, both low-performing districts, administered more tests overall than Hillside or Mannequin, the two high-performing districts. Figure 5 below shows the districts ranked in order of total tests administered in grades K-12: Sunset (221 tests), Overton (209 tests), Hillside (102 tests), and Mannequin (79 tests).

Although Sunset and Overton were fairly close in the number of tests administered, Sunset was an extreme outlier in terms of time spent testing with an average of 48.9-49.2 hours per grade on state and district tests. According to recent research from the Council of Great City Schools, Sunset's average testing time per grade is greater than 66 of the nation's largest school districts in all grades but kindergarten and first grade.[14] After Sunset, the other districts in order of most-to-least hours testing are Overton (15.6-16.2 hours), Hillside (13-13.3 hours), and Mannequin (11.9-13.9 hours). Overton, the district with the second-highest testing time, averaged about 30 hours less time per grade per year than Sunset.

Sunset's abnormally high testing time is the direct result of its lengthy nine-weeks exams given in each core subject in each grade 2-12. While all Mississippi districts mandate core subject nine-weeks testing in many grades, Sunset's approach makes it exceptional. We acknowledge that one quirk of our methodology is that nine-weeks tests created by the district appear in our results while teacher-created nine-weeks tests do not. This alone, however, does not explain

FIGURE 5. Number of District-Mandated Tests, K-12



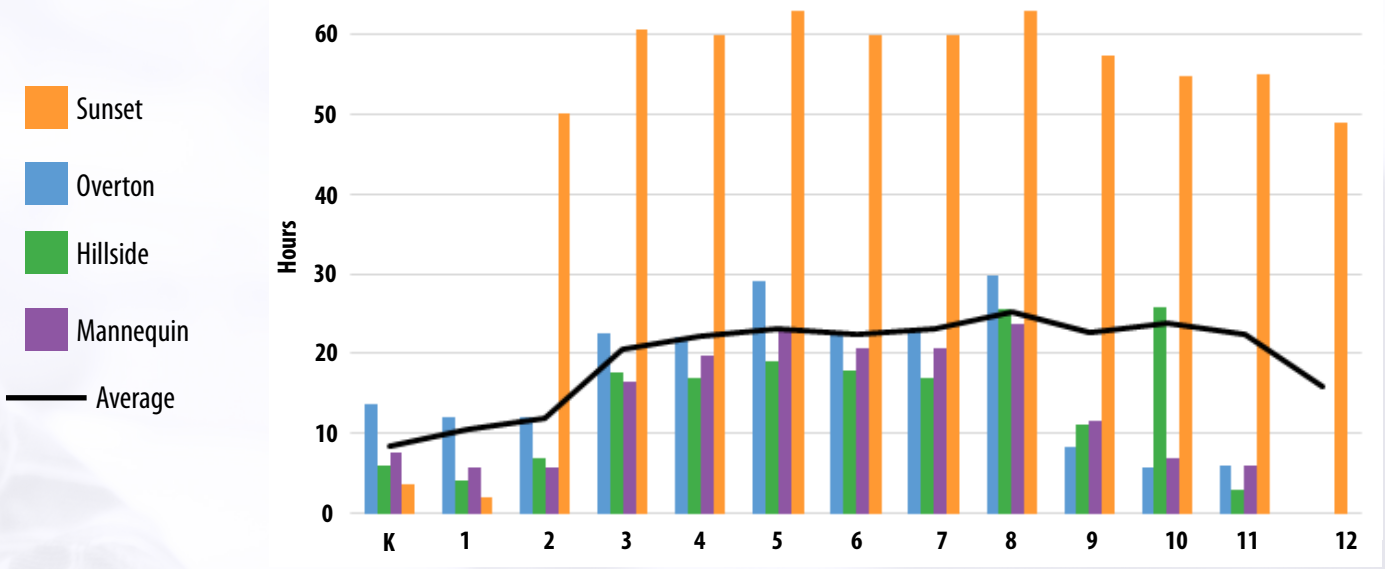
[14] Hart, Ray, Michael Casserly, Renata Uzzell, Moses Palacios, Amanda Corcoran, and Liz Spurgeon. 2015. *Student Testing in America's Great City Schools: An Inventory and Preliminary Analysis*. Research Report, Washington, D.C.: Council of Great City Schools. Accessed June 23, 2016. http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Testing%20Report.pdf.

COMPARATIVE FINDINGS, CONTINUED

testing time in Sunset: the driving factor is that Sunset's tests are extremely long, even in the earliest grades. Each test is three hours! In other words, Sunset's students take *four* three-hour tests each quarter from second grade to twelfth grade. That's a total of 12 hours of district testing each quarter and 48 hours of district testing per year just from these assessments. The fact that Sunset teachers and administrators told us that they believe many teachers do not use the data generated from these long tests only adds insult to injury. If Sunset reduced its nine-weeks tests to just one hour per subject for the first three nine weeks and 90 minutes per subject for the last nine weeks, it could cut its district testing time by two-thirds, bringing it more in line with the other districts that we studied. The need for data support is discussed in Finding 3.

Figure 6 below shows the total time spent completing tests by grade in each school district in 2014-2015.[15] The black line shows the national average among the nation's 66 largest school districts from the recent Council of Great City Schools' study.

FIGURE 6. Time Spent Completing Tests in Each Grade by District



[15] In Figure 6 only, for purposes of visualizing this data, we report expected completion time as the halfway point of any given range. For example, a test taking 50-75 minutes has been reported as 62.5 minutes in Figure 6. Remember that this data represents the minimum amount of instructional time lost for the purposes of standardized assessment. Actual time lost may be much greater, depending on the efficiency of test administration.

COMPARATIVE FINDINGS, CONTINUED

COMPARATIVE FINDING 3: Teachers in high-performing districts may receive more support for data review and analysis. In our focus groups, teachers in each of our districts discussed using available data to inform their planning or identify which objectives were mastered by students. Although we did not specifically ask about data analysis structures as part of our focus group protocol, teachers' comments about formal data analysis—or the lack thereof—were revealing. All teachers described wanting more support in reviewing and analyzing data for decision-making, but only teachers in high-performing districts described formal procedures for reviewing data from at least one of their district-mandated tests. These procedures gave teachers dedicated time to review data and discuss it with administrators, either one-on-one or in a group of other teachers. Teachers felt that these meetings provided valuable learning opportunities, a venue to make strategic instructional decisions, and an important internal accountability mechanism.

Teachers in the low-performing districts did not discuss any specific data analysis procedures, indicating that they did not receive the same level of support as teachers in high-performing districts. Teachers in Overton were vocal about needing more support to better utilize the data features in their vendor testing products and to prioritize which data to use for instructional decisions. Sunset teachers, as well as the district's instructional strategists, openly questioned whether all teachers used the data provided by various Sunset tests to drive instruction, indicating there was no process to ensure the use of data.

COMPARATIVE FINDING 4: 1-to-1 technology did not reduce test completion hours, although it may have reduced testing-related disruptions. One of our initial research questions sought to investigate the impact that factors like available technology would have on testing time. With the surge of computer-based assessments, we wanted to know whether districts with 1-to-1 access—those with one device for every child—were at an advantage in reducing testing time. We also wanted to discover whether these districts approached testing differently due to the prevalence of technology. In our sample, Hillside students in grades K-12 and Sunset students in grades 6-12 had 1-to-1 access to iPads. The other two districts did not provide 1-to-1 access to technology for students.

We found that technology access did not have an impact on the amount of time students spent completing tests. Students in Sunset, with partial 1-to-1, spent the most time completing tests while students in Mannequin, which tied with Overton for the lowest rate of testing-capable devices per student, spent the least amount of time completing tests. Technology also did not change districts' testing strategies as both Hillside and Sunset also mandated at least some paper-and-pencil tests.

The availability of testing-capable devices did appear to impact learning, just not in terms of test completion hours. Teachers in Mannequin were most outspoken about this problem. For state testing, Mannequin students had to rotate through a small number of computer labs over a longer period of time, tying up school resources for instruction. Mannequin also had this problem with its progress-monitoring exams, all of which are computerized. In order to get whole grade levels into the computer labs to take progress-monitoring tests three times a year, the labs were often booked for several days, even weeks, at a time. Teachers in Mannequin felt that computer labs were not available for regular classroom instruction between district testing and state testing. Overton also listed "scheduling students into [computer] labs" as a factor prolonging testing-related disruptions.

Teachers in all districts wanted **more support** for **data analysis**. Only teachers in high-performing districts in our sample received any support.

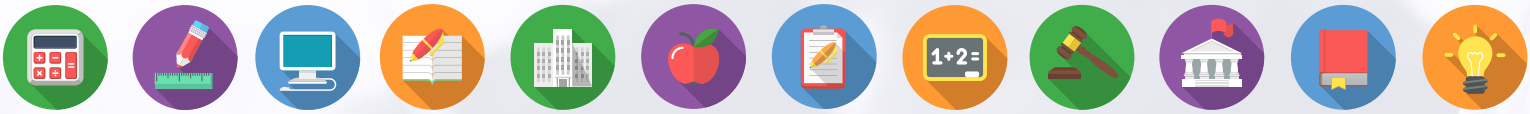
COMPARATIVE FINDINGS, CONTINUED

COMPARATIVE FINDING 5: Low-performing schools prioritized test prep over content instruction for at least 25% of their instructional year. We were alarmed to find that teachers in both of the low-performing school districts in our sample stop teaching content at the end of the third nine weeks and begin daily, intensive test prep. In Sunset, this practice is considered an official policy in the district. In Overton, teachers stated that they "should not be starting on new standards" in April, but it was less clear whether they were required to stop teaching new content or if teachers simply universally accepted this practice. Regardless of the reason, the results show that students in the low-performing districts in our sample are losing at least a quarter of the instructional year ***by design***.

The issue appeared even more severe in Sunset: all of the teachers in Sunset are attempting to cram an entire year's worth of learning in approximately ***two-thirds*** the time. Due to Sunset's practice of stopping instruction during district nine-weeks testing, teachers in K-8 lose not only the last quarter but also the three full weeks that are devoted to nine-weeks testing up to that point. High school teachers, whose courses are on a 4X4 block schedule, face a similar challenge—out of an 18-week semester, they must attempt to cover their entire course in 13.5 weeks. In that time, they also take district benchmark tests every quarter, which is a loss of at least another week of instructional time within the semester.

In Review:

1. Districts have **wide autonomy** in how they use testing, resulting in **very different student experiences** across districts.
2. Low-performing districts in our sample **administered more tests** and **spent more time** testing than high-performing districts.
3. Teachers in high-performing districts may **receive more support for data review** and **analysis**.
4. **1-to-1 technology** did not reduce test completion hours, although it may have reduced testing-related disruptions.
5. Low-performing schools **prioritized test prep over content instruction** for at least 25% of their instructional year.



RECOMMENDATIONS

School districts, MDE, and the legislature all have a role in improving the use of standardized testing in Mississippi. In this section, we identify three broad recommendations followed by specific actions for each of these policymakers.

RECOMMENDATION 1:

Mississippi should increase transparency about testing in public schools.

After conducting this study, we still have many questions about how widespread some of the testing practices we uncovered are, especially those we found detrimental. Statewide, there is no source for good information about testing in each district and no requirement that anyone report this information. This lack of transparency causes problems not only for researchers but also for parents and legislators.

Teachers in our focus groups repeatedly told us that parents struggle to differentiate between district and state tests and to understand the purpose of each test. This confusion adds to parents' belief that children are being over-tested, particularly in the younger grades. Parents are not the only ones who struggle to understand the purposes behind each assessment: legislators do not always appear to know the difference between state and district tests. When legislators act without a solid grasp of all the facts, the consequences can be enormous for every public school child.

RECOMMENDATION 2:

Mississippi must put testing back in its appropriate place in education.

When both state and district tests are used appropriately, students benefit.

State-mandated tests primarily serve an accountability function for policymakers and the public: they let us know whether students are learning what we expect by the end of every grade, and they allow us to compare students in different districts across the state. State tests also serve an accountability function for parents. Parents can use state assessments as an independent evaluation of their children's progress in school, and they can compare state test results to grades to determine if their children have truly mastered the material and if the school is serving their children well. Teachers, too, can use state assessments to determine if they need to alter their practices for the coming year.

District-mandated tests primarily serve an instructional function, although they may also serve an internal accountability function in some districts. Teachers and district leaders use district tests to understand how students are progressing throughout the school year so that they can make adjustments to lesson plans and strategies. Parents can also use district-mandated tests to monitor their students and know if their children are veering off course before it is too late.

Our research shows that many children experience testing as more of a hindrance than a help. It points to different solutions for state and district testing.

RECOMMENDATIONS, CONTINUED

Due to federal requirements, eliminating annual state testing is not an option. We want to be very clear: we would not recommend eliminating annual state testing even if the federal government did not require it. Mississippi respondents to a recent scientific statewide poll agree: a majority (56.7%) "agreed" or "strongly agreed" with annual state testing, with the largest proportion of respondents strongly agreeing (37.9%).[16] Teachers in our focus groups mostly agreed as well, even if annual state testing has downsides.

Instead, Mississippi should improve state testing by making the tests more valuable to teachers for instruction and by carefully weighing whether changes to the testing program will be worth the pain of transition. For district testing, districts must ensure that the benefits of testing outweigh costs. The research reflected in this report shows that districts are wielding their autonomy in inconsistent ways, some of which are not helpful to student progress.

RECOMMENDATION 3:

End the overreliance on test prep.

The most alarming discovery we made in the course of this research was that the low-performing districts we studied purposefully designed their pacing guides to end after three quarters of the school year. Neither of our A-rated districts engaged in this practice, though all teachers discussed test prep. Since our research did not include surveying all D- or F-rated districts, we do not know how widespread this practice is, although several education leaders we spoke with shared anecdotes that indicate this practice may be very common in low-performing school districts.

Shortening the instructional year is extremely counterproductive because students are unlikely to grasp all of the new concepts in the time allotted, and even the best students are unlikely to understand the new concepts at any depth. Furthermore, students who struggle academically need more instructional time, not less. The hard truth is that review of content that students did not have the time to learn in the first place will have little effect.

We understand the pressure that low-performing districts must feel and the impulse to lean heavily on test prep in these circumstances. High-performing districts may also struggle with this at a classroom level, even if quarter-long test prep is against official policy and pacing guides. Regardless of whom engages in this practice, spending 25% of the school year on test prep is damaging to students' academic progress. Multiple research studies show deep content learning is more effective than test prep at improving student achievement.[17] Ironically, in an effort to propel students forward, districts emphasizing test prep for a quarter may actually be denying students one of the most critical supports for learning—time. We cannot stress enough how deleterious this practice is.

---

[16] Canter, Rachel, Angela Bass, and Searcy Milam Morgan. 2017. *Mississippi Voices: Public Perception of Pre-K-12 Education in Mississippi*. Research Report, Jackson, MS: Mississippi First. Accessed February 9, 2017. http://1iq0332x28t34od07uajkv11.wpengine.netdna-cdn.com/wp-content/uploads/2015/11/Mississippi-Voices.pdf.

[17] The Meadows Center for Preventing Educational Risk. 2017. *10 Key Policies and Practices for Assessment in Schools.* Research Summary, Austin, TX: University of Texas-Austin. Accessed November 15, 2017. https://www.meadowscenter.org/files/resources/10Key_Assessment_Web.pdf.

WHAT DISTRICTS CAN DO

Districts experience the costs and benefits of testing the most directly of all policymakers. They also have the power to decide how most testing occurs in classrooms.

TO INCREASE TRANSPARENCY

School districts should publish a table of all their standardized tests, with state testing and district testing clearly delineated. We do not recommend mandating this table through MDE or the legislature because we hesitate to add an additional legal or regulatory paperwork burden on school districts. Instead, we believe that districts that voluntarily take this step will find that it is a valuable tool for both district strategy planning and parent communication, and well worth the time it takes to construct. Districts can use a variety of formats for this table, but we have suggested one in Appendix D.

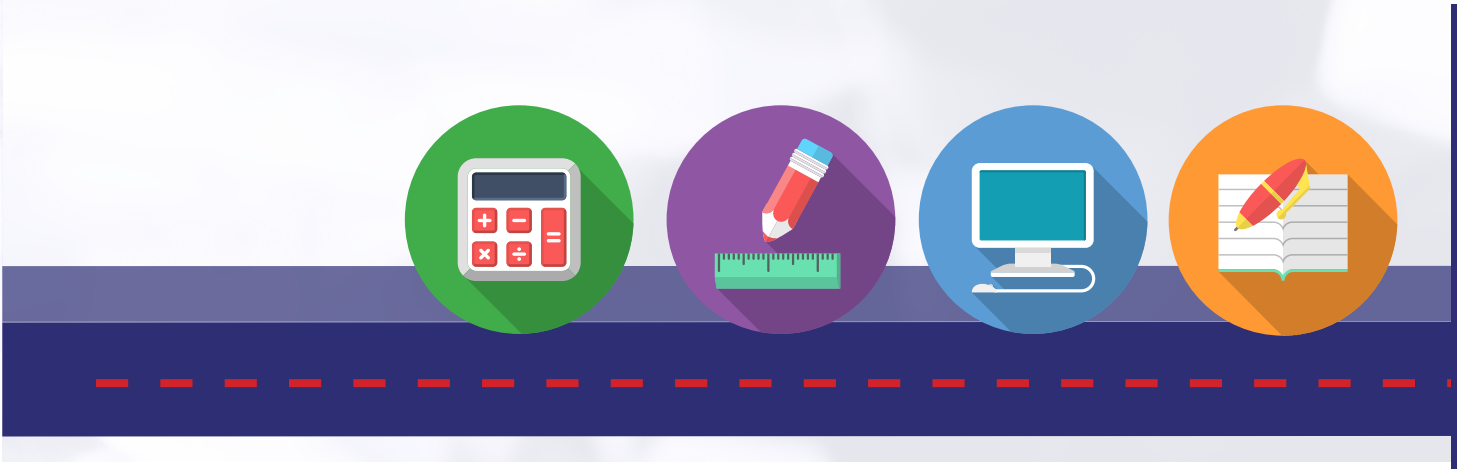
School districts should host a session on testing at "back-to-school" night, their "parent university," or other regular parent meeting. Most districts have long-established events for explaining school policies to parents, but many do not use these opportunities to discuss testing with parents. Mannequin began talking about testing in its parent university and found that parent angst over testing improved after parents had better knowledge about what tests measure and how they are used.

TO PUT TESTING IN ITS APPROPRIATE PLACE

Support teacher review of state testing data through formalized planning procedures in the new school year. Teachers' biggest gripe about state testing is that the data are not helpful in improving instruction, since most teachers get new students year-to-year and end-of-year results do not help teachers improve learning for students they no longer teach. If teachers cannot use the data to remediate the students who took the test, teachers' next best use of the data would be to inform their planning for teaching the same objectives to new students. Some teachers do a deep analysis of their prior-year data in the week before school starts, which is generally all the time they have for this task. Many teachers, though, find last year's test data too overwhelming to analyze beyond a cursory review, due to the hustle to get everything in order for a new year. With a little organization, districts can make state test data useful to teachers in their instructional planning by taking some of the legwork out of analysis, scheduling mandatory review time, and facilitating data conversations. A good model for this in our sample was Hillside's: teachers found one-on-one data meetings with coaches to review end-of-year results valuable even though they no longer taught the same students. Districts can also provide students' new teachers with detailed reports of their prior-year test scores as a diagnostic aid.

Conduct an audit of all standardized testing occurring throughout the district to collect information on the types of tests administered, how much time they require, and the purposes they serve. Achieve, a nonprofit organization dedicated to working with states to improve assessments, provides a free tool for districts to conduct comprehensive testing audits.[18] After the audit, administrators should work with teachers to determine the minimum testing necessary to serve essential instructional and accountability purposes. Administrators and teachers should ensure that every standardized test retained is high quality and supported by structures and routines so that assessment results can be used to help students.

[18] For more information about Achieve, visit their website at www.achieve.org.



WHAT DISTRICTS CAN DO

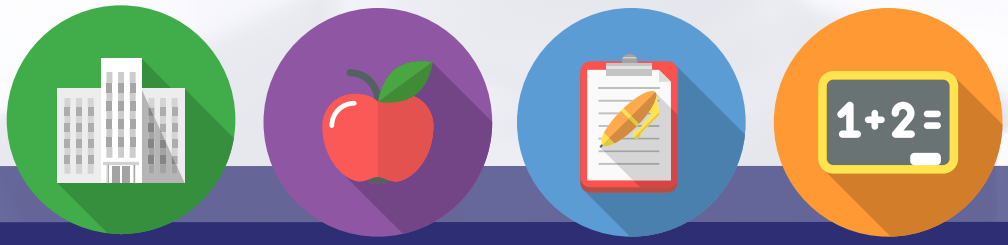
Increase support for rigorous teacher-created tests. District staff can take a strong role in facilitating quality teacher-created tests that accurately and consistently track progress, thus eliminating the need for additional tests. Ongoing workshops and coaching on quality assessment practices would increase teacher skill for every formal assessment a teacher creates and administers. The district should also be mindful of how teachers of different experience levels need more or less support. Several teachers requested help in writing tests more aligned to state tests in format and rigor. One of the teachers interviewed for this project stated that as a newer teacher, she appreciated being given quality test questions from the district assessment because she does not have years of assessments to select questions from and needs examples of good tests. More experienced teachers may simply need to have their test questions reviewed for rigor and/or alignment.

Only administer a test if there is a clear plan for using the data. A district-mandated test should always support instructional decision-making. Districts assume that if a test is mandated, teachers will make use of the data without further action on the part of the district, but our research shows this is often a faulty assumption. Teachers in every district we sampled relied upon the district to organize data meetings and lead data discussions. Even with sophisticated data reports from vendor-created assessments, teachers still sought one-on-one and group data meetings in which they could make sense of all the data they receive. Due to the amount of data they are receiving from multiple assessments, teachers sometimes needed help prioritizing what they should focus on. Teachers in high-performing districts may be more likely to receive this type of support, but these teachers asked for more support anyway. Teachers in low-performing districts in our sample also sought this help and appeared less likely to receive it, a huge opportunity for improvement.

TO END THE OVERRELIANCE ON TEST PREP

Rewrite pacing guides to protect instructional time. As part of their annual process to review and revise pacing guides, districts should pay special attention to how their pacing guides balance learning new content with review and remediation. Districts with pacing guides that end after three nine weeks should give teachers and students more time to learn content well on the front end. We suggest that districts begin revising pacing guides in ELA and math in grades 3-8 before moving to pacing guides in other grades and subjects.

Adopt district policies to protect instructional time. District leaders, including the school board, are in the best position to prevent test prep from overwhelming school calendars. District policies governing pacing guides, review sessions, length of the school day and year, and district nine-weeks testing, to name a few, can deeply affect how much instructional time students receive. Districts should review their current policies to see which ones may be adversely impacting instructional time and make adjustments. They should also determine whether they need new policies to clarify expectations for teachers and administrators in relation to instructional time.



WHAT THE MISSISSIPPI DEPARTMENT OF EDUCATION CAN DO

MDE often receives the brunt of criticism about testing, but its role in improving testing is mostly a supporting one. MDE should focus its efforts on helping stakeholders, particularly parents, teachers, and school district leaders, better understand and utilize testing for its intended purpose—gauging student progress.

TO INCREASE TRANSPARENCY

✓ **MDE should create an easily accessible parent testing guide about state tests for use on its website and for distribution to districts.** Currently, state testing information on MDE's website is geared toward school districts rather than parents. There are specific parent guides for MAP, but these provide general information rather than testing times per grade. The parent guide we propose can be very short—a summary telling parents what to expect at each grade level and a table of testing times. MDE has recently produced a number of high-quality guides for stakeholder audiences, so this guide would be in line with that strategy.

TO PUT TESTING IN ITS APPROPRIATE PLACE

✓ **Press state test vendors on shortening turnaround time between test completion and score reporting.** Teachers expressed two desires for state testing that have historically been in tension: they want state tests administered as close to the end of the year as possible while at the same time wanting the results before they move on to the next school year. For 2017-2018, the state testing window is from mid-April to mid-May, and state test results are not expected until July. Teachers have usually left for the year by the beginning of June and do not return until the beginning of August. This leaves little time for teachers to analyze state test data before they must begin teaching a new set of students. Ideally, all state tests would have the same turnaround time as the MKAS[2], so that teachers could review data with their administration prior to leaving for the summer.

✓ **Consider test completion time when making state test decisions.** PARCC had an ambitious goal—design high-quality, "next generation" standardized tests measuring new state standards across dozens of states. As part of this next generation plan, PARCC prominently featured performance-based tasks completed during a separate test administration. While performance tasks are now the norm nationwide, the separate administration doubled not only the number of parts to the state assessment but also the time that students spent testing. Teachers we interviewed believe the additional time outweighed the value of the new information tested, a lesson to keep in mind for future testing decisions.

✓ **Over communicate big changes to the state testing program directly to teachers, and slowly implement those changes whenever possible.** Teachers repeatedly commented that the rapid change from MCT2 to PARCC to MAP (now MAAP!) was almost too much to bear. Having three tests in three years was no one's optimal scenario, including MDE, which got stuck in the unenviable position of having to scramble for a new test after the PARCC contract was delayed during the test's planned second year. From an educational perspective, teachers in our sample understood the need to replace the MCT2, not only because the standards changed but also because the MCT2 was not a good indicator of student knowledge. Nonetheless, PARCC ushered in several changes all at once, including testing entirely online, a performance-based assessment separate from the end-of-year test, new question types, and enhanced rigor. Because PARCC itself was new, the state did not have enough lead time to fully communicate the impact of all of these changes to teachers. Teachers felt that their own lack of knowledge about the test meant that they could not adequately prepare their

students for all of the changes. After testing, the long delay in receiving the data made teachers feel like all of the frustration and anxiety did not even have a good purpose. Teachers were happy to see PARCC go, but they felt similar uncertainty and frustration about MAP. The circumstances leading to the rapid succession of state tests will hopefully not arise again any time soon, but one takeaway is that MDE could reduce feelings of anxiety caused during the roll-out of new tests by overcommunicating directly to teachers, especially by disseminating sample test items demonstrating the changes as early as possible. MDE should also slowly transition through major test upgrades, if at all possible.

Build technical assistance capacity to help districts audit and redesign their testing practices. The school districts that we audited had never had a third party examine their testing practices to recommend improvements. What we found is that their testing practices had evolved over time, generally without specific strategic conversations as to why the district was employing certain tests or methods. In some cases, school-level administrators or even grade-level chairs had autonomy to choose assessment products, leading to an incoherent system districtwide that changed as personnel in decision-making roles changed, rather than by strategic design. This incoherence was especially acute in the lower-rated districts we studied. In these circumstances, both students and teachers experience testing as more of a burden than a help. MDE is well positioned to support districts in rethinking their testing programs and improving their effectiveness.

Apply for funds through the State Assessment Grant Program to conduct testing audits, if and when such funds become available. Under the *Every Student Succeeds Act* (ESSA), the U.S. Department of Education (Ed) has the ability to earmark funds from the State Assessment Grant Program for testing audits. If Ed chooses to do this, state grants would be up to $1.5 million, with at least 20% directed to districts. Thus far, the feds have not chosen to make these grants available.[19] In the meantime, states can re-allocate funds they already receive by formula from the State Assessment Grant Program for this purpose.

Encourage teacher and administer preparation programs to develop and implement courses on understanding, designing, developing, and analyzing assessments. Far and away, teachers preferred their own tests for tracking student mastery of learning standards. From the perspective of teachers, their own tests were the most relevant and even the most rigorous. In a perfect world, teacher-created assessments—whether created jointly by grade-level teams or by individual teachers—would be all schools and districts need to evaluate student progress. In reality, teachers have varying levels of knowledge and skill in developing tests and analyzing results. This can lead to a muddied picture of student learning from classroom to classroom.

Research from the National Council on Teacher Quality shows that despite the heavy importance of data and assessment in modern education, teachers receive little to no instruction in teacher prep programs on how to develop a quality assessment or how to analyze data from one.[20] The logical conclusion is that any knowledge teachers have of how to write or use a rigorous, valid, and beneficial test comes from professional development teachers receive while in service. It makes far more sense to strengthen teacher assessment skills during teacher preparation, rather than relying on districts to fill in this knowledge.

TO END THE OVERRELIANCE ON TEST PREP

Publish model pacing guides for all grade levels, beginning with grades 3-8. Adults often learn best by seeing examples of effective practices. While MDE could invest staff resources in developing model pacing guides, we believe it may be more efficient to collect and publish exemplary pacing guides from high-performing school districts. MDE could also create a database of pacing guides that correspond to curricula in use by Mississippi districts so that districts using the same math curricula in fourth grade, for instance, could compare pacing guides and adjust when another district has a better model.

Provide technical assistance to districts in rewriting pacing guides. Even if districts adopt new, exemplary pacing guides created by their high-performing peers, the instructional planning skills required to develop pacing guides are necessary to every district. Districts with unique curricula in some or all subjects will also need support if there is no matching exemplary pacing guide.

[19] Klein, Alyson. 2018. "Answer Your ESSA Questions: What's Going on with Testing Audits?" Politics K-12, *Education Week*, January 3. http://blogs.edweek.org/edweek/campaign-k-12/2018/01/answering_essa_questions_testing_audits.html.

[20] Greenberg, Julie, and Kate Walsh. 2012. *What Teacher Preparation Programs Teach about K-12 Assessment: A Review.* Research Report, Washington, D.C.: National Council on Teacher Quality. Accessed June 23, 2016. http://www.nctq.org/dmsView/What_Teacher_Prep_Programs_Teach_K-12_Assessment_NCTQ_Report.

WHAT THE LEGISLATURE CAN DO

Legislators play a unique role in the education policy landscape. They can make change not only by passing legislation, but also by investigating problems and educating the public about issues. Legislators are most helpful to educators when they play to their strengths by focusing on setting broad policy goals and providing the resources to achieve those goals, rather than veering into functions best left to an executive agency or a school district.

TO INCREASE TRANSPARENCY

Avoid adding to the confusion with overblown rhetoric about testing. Politics often does not lend itself to measured commentary on complex and arcane topics like testing. Nonetheless, legislators can create more public understanding by avoiding blanket statements about testing that no one has yet proven.

Investigate testing practices in the districts that they represent. Knowing about the unique practices of each school district will give legislators context for parental or educator complaints and may suggest local solutions that can be more effective than broad or blunt state policy changes. After carefully examining each districts' data in this study, we recommended different solutions to each. For Sunset, for example, we recommended reducing the length of each district benchmark test. If we had suggested a state policy requiring every district to reduce nine-weeks testing time, however, we would not have solved Overton's biggest problem, which was having duplicative progress-monitoring tests in every grade. Legislators should understand the nuance of the testing issue through concrete examples in order to make effective policy.

TO PUT TESTING IN ITS APPROPRIATE PLACE

Leave vendor testing decisions to MDE. For the last few years, at least one legislator has filed a bill each session to dictate the vendor for state testing. The 2018 efforts by Representative Tom Miles and Representative Gary Chism died on committee day without a hearing. Although both bills were widely characterized as replacing only the four high school subject-area tests with the ACT college entrance exam, Rep. Miles's bill would have also replaced ELA and math state tests in grades 3-8 with the ACT Aspire products. In February 2017, the U.S. Department of Education questioned whether Alabama, one of three states in the country which used ACT Aspire for state testing, could show that the ACT Aspire aligned to Alabama's state standards, which are very similar to Mississippi's.[21] Alabama voted to end its ACT Aspire contract in June 2017, citing problems with ACT Aspire as a vendor.[22] Currently, Arkansas and Wisconsin are the only states using ACT Aspire for accountability. Both states use the test for grades 3-10, but neither state has yet passed peer review for its use of the ACT Aspire.

[21] U.S. Department of Education, letter to the Honorable Michael Sentence, Alabama State Department of Education, January 6, 2017. Accessed March 27, 2018. https://www2.ed.gov/admins/lead/account/nclbfinalassess/al6.pdf.

[22] Aubuchon, Alex, and the Associated Press. 2017. "Alabama Drops ACT Aspire Test." *Alabama Public Radio*, June 22. http://apr.org/post/alabama-drops-act-aspiretest#stream/0.

The ACT college entrance exam has also not yet been approved through peer review for primary accountability use in any state. In early 2017, the U.S. Department of Education questioned whether the ACT college entrance exam was aligned to high school standards in Wisconsin and Wyoming.[23] Both Wisconsin and Wyoming have similar ELA and math standards to Mississippi's. National alignment studies show clear problems with using the ACT in this way.[24] For example, January 2018 research comparing the ACT and Florida's state standards, which are also similar to Mississippi's, revealed that the ACT was not well aligned.[25] If Mississippi were to adopt the ACT as its high school exam or if Mississippi were to allow districts the flexibility to choose ACT as the exam, Mississippi would have to produce independent technical documents demonstrating ACT's alignment to our state standards and its suitability for use as an accountability test.[26] While no state has been approved to use the ACT for accountability, North Dakota applied for a waiver of the peer review process in February 2018 in order to allow districts to choose ACT over North Dakota's high school exam this year. That waiver was granted on March 5, 2018, but North Dakota must complete peer review prior to 2019 in order to continue this practice.[27] Even if North Dakota successfully completes peer review, it is no guarantee that Mississippi will.

Not only did Rep. Chism's bill replace current high school tests with the ACT, it also explicitly made a minimum score on the ACT a requirement for graduation. Setting and requiring a cut-score on a college-readiness exam as a basis for high school graduation when not all students intend to pursue college is inherently problematic. ACT has further stated that its exam is not intended for use as a high school exit exam.[28] A more appropriate use of the ACT is how Mississippi is using it now—to demonstrate college readiness.

Lastly, the issue of adopting a test via legislation is problematic beyond whether or not the ACT or ACT Aspire are specifically appropriate. Legislators are often not in the best position to understand and weigh all of the complex factors that must be part of the decision about which test to adopt. With the ACT, for example, there are a number of key technical issues to consider, including the ACT's inability to differentiate among student scores lower than a 12 and its challenges with testing children with disabilities who require accommodations. Furthermore, sole-source contracting via legislation opens a lot of ethical questions. Contracting via a request for proposals is a job best left to a state agency like MDE, with appropriate oversight.

Appropriate funds allowing MDE to help districts audit and redesign their testing practices. MDE is frequently called upon to assume new duties or re-prioritize their work based on changing conditions, but these requests rarely come with the resources to support the new work. While MDE can likely re-allocate some funding to support this technical assistance, the legislature should boost MDE's budget for this purpose to prevent the Department from having to reduce its effectiveness in another area.

TO END THE OVER RELIANCE ON TEST PREP

Create a taskforce to study inequities in access to instructional time across districts. Districts engaged in over-testing or long test-prep periods rob their students of instructional time best dedicated to learning new material well in the first place. It is not at all clear to us from this research that regulating testing time as Rep. Guice sought to do in his 2017 bill would be the best solution, as the appropriate amount of time for testing is context dependent. We recommend that the legislature convene a blue-ribbon panel to determine the scope and causes of the problem before settling on a solution.

[23] Gewertz, Catherine. 2017. "U.S. Ed. Dept. Defers Approval of ACT for Accountability in Wyoming, Wisconsin." *Education Week*, January 20. http://blogs.edweek.org/edweek/high_school_and_beyond/2017/01/us_ed_dept_defers_approval_of_act_for_accountability_in_wyoming_wisconsin.html

[24] Gewertz, Catherine. 2018. "Don't Use SAT and ACT as Your High School Tests, Study Urges." *Education Week*, March 13. https://www.edweek.org/ew/articles/2018/03/13/dont-use-sat-and-act-as-your.html.

[25] Roeber, Ed, John Olson, and Barry Topol. 2018. *Feasibility of the Use of the ACT and SAT in Lieu of Florida Statewide Assessments: Volume 1: Final Report*. Research Report, Tallahassee, Florida: Assessment Solutions Group. Accessed March 27, 2018. http://www.trbas.com/media/media/acrobat/2018-01/70109708365300-05065523.pdf.

[26] U.S. Department of Education, letter to State Assessment Directors and State Title I Directors, May 15, 2017. Accessed March 27, 2018. https://www2.ed.gov/admins/lead/account/saa/locallyselected72117.pdf.

[27] U.S. Department of Education, letter to the Honorable Kirsten Baesler, North Dakota Department of Public Instruction, March 5, 2018. Accessed March 28, 2018. https://www2.ed.gov/admins/lead/account/nclbfinalassess/nd5.pdf.

[28] Harris, Bracey. 2018. "Here's Why Mississippi Won't Swap Exit Exams for the ACT." *Clarion-Ledger*, January 31. https://www.clarionledger.com/story/news/politics/2018/01/31/mississippi-keeps-exit-exams/1082389001/.

APPENDIX A: RESEARCH TOOLS

For this project, we used adapted research tools from Achieve, a national education policy organization that conducts nationwide research on standards, assessments, and accountability. Achieve developed the Student Assessment Inventory to support school districts in determining the appropriate balance of testing necessary to serve essential diagnostic, instructional, and accountability purposes. Mississippi First modified the Student Assessment Inventory to capture more information about how assessments impact the whole school environment. The original tool from Achieve can be found here: http://achieve.org/assessmentinventory. In addition, we developed a Technology Inventory to assess the technological capabilities in each district. This tool was developed with guidance from some technology experts in Mississippi school districts. Finally, we used a focus group guide from Achieve to create our teacher focus group protocol. Below are the tools we used.

Student Assessment Inventory for Districts (Adapted from Achieve)			
Name of Assessment	Test 1	Test 2	Test 3
Entity requiring assessment			
Grade(s) Tested			
Course(s) or subjects tested			
Which students are eligible or required to take assessment?			
Type of assessment *Summative; interim/benchmark; formative; diagnostic*			
Number of years assessment has been administered in the district			
To which content standards is the assessment aligned? *(source of alignment verification)*			
Intended purpose(s) of the assessment			
Intended use(s) of the assessment			
Users of the assessment			
Do users of the assessment use it for its intended use(s)?			
To what degree do users of the assessment find it useful or not useful? *1 – not useful* *2 – somewhat useful* *3 – useful* *4 – very useful* Explain why.			
Type of administration			
Item type(s)			
Accommodations			
Expected test administration time			
Common factors that impact test administration time			

APPENDIX A: RESEARCH TOOLS, CONTINUED

Student Assessment Inventory for Districts (Adapted from Achieve)			
Name of Assessment	Test 1	Test 2	Test 3
Testing window			
Test frequency			
Number of days required to test all student testers			
Disruptions to non-testing students during testing days			
Disruptions to non-tested subjects during testing days			
Number of days dedicated to test review			
Other disruptions to school day not captured			
Time between test administration and results to users			
Vendor			
Contract expiration date			
Entity that holds contract			
Annual cost *(total and per student)*			
Funding sources			

APPENDIX A: RESEARCH TOOLS, CONTINUED

Technology Inventory for Districts (Created by Mississippi First)	
Name of School	School 1
Number and types of devices available to students for assessments (do not include devices only available to teachers or staff)	
Vendor(s)	
Cost of devices *(total and per student)*	
Funding sources	
How old are your devices?	
What is the life expectancy of your devices?	
Number of computer labs	
How many technicians do you have on staff? How many instructional technologists do you have on your staff? What are the responsibilities of each?	
Internet bandwidth into district	
Internet bandwidth between buildings (school to school and between buildings on campus)	
Do you have a direct connection to the internet or does your internet connect back to the central office?	
Are there wireless access points in every classroom?	
Are there enough access points for the number of devices at the school?	
Can all devices be on the network at the same time?	
Do all devices have updated virus protection software?	
What is your process for keeping devices protected from viruses?	
Do any of your districtwide or schoolwide tests require the installment of software? Which ones?	
What is your process for installing software and keeping it up to date?	
Describe the training that users (admin, teachers, students) of the software get. Is it required?	
If there is an issue with the assessment or the software, who can fix it? What is the response time?	
If there is a technical issue with a device, who can fix it? What is the response time?	
Do you track response time to tickets when a device is broken?	
Who controls student passwords and PINs for assessment software? What is the response time for resetting forgotten passwords and PINs?	
Describe how technology is factored into the planning process for assessments.	
What student testing accommodations are you able to provide through technology? Who administers accommodations?	
Common factors with technology that impact test administration time?	

APPENDIX A: RESEARCH TOOLS

Focus Group Protocol (Adapted from Achieve)

Introduction of Moderator and Focus Group Guidelines

- My name is_____, and I'll be moderating today. Thanks for agreeing to talk with me. This conversation will be extremely casual, and I want you to feel as comfortable as possible. I'm here on behalf of the organization that I work for, Mississippi First, to gather information from you. I will meet with three groups of teachers in your district today.
- Mississippi First is a nonprofit organization that specializes in education policy and advocacy. We advocate for policies that are best for Mississippi kids. We are interested in conducting a research project to gain a better understanding of the time devoted to testing in districts. We will use our research to make informed policy recommendations to education leaders and policymakers in the state, including district superintendents, state education officials, and legislators. All of our findings will be published in a report that will be available to the public. The names of districts or people will NOT be published in public reports.
- We will be talking about your thoughts and experiences with testing at your schools. This will help us and your district leaders learn about what testing is like at the classroom level, and it will help provide an understanding of the quality and use of assessments given in the district.
- One person should speak at a time.
- There are no "wrong" answers to any of these questions. We are interested in hearing your perspectives as teachers.
- We value confidentiality. We will not associate any feedback that comes out of this focus group with a particular individual.
- Everyone will have a chance to speak. If you have not had an opportunity to provide your perspective, I may call on you.
- Are there any questions or additional norms you would like to add?

Disclosure of Note-Taking

- I will be taking notes to make sure we get all of your feedback. I will not associate feedback with names.

Parking Lot

- There is a "parking lot" chart where we can put our ideas or thoughts that come up in our discussion that are important but may not be related to the purpose of this group. We want to capture those important thoughts, but we also want to keep focused on the purpose of our meeting. These ideas or thoughts will be shared with appropriate individuals following the meeting.

Communicating Results of Focus Groups

- The results of the focus groups will be summarized, and you will receive a summary by December 31. If we missed any key points you raised during this conversation, please let us know.

Focus Group Questions

1. What grades and subjects do you teach? How long have you worked in the district?
2. How have you seen testing change in the last five years?
3. Do you think students in this district are given too many assessments, not enough assessments, or about the right number of assessments throughout the school year? Why?
4. What are you hearing about assessment from other teachers? What are you hearing from parents? What are you hearing from students?

Teachers	Parents	Students

APPENDIX A: RESEARCH TOOLS

Focus Group Protocol, continued

5. What are some examples of district assessments that have been helpful to your work, such as informing instructional practice, diagnosing student needs or predicting a student's later performance? How have they been helpful? *(Moderator prompt: strong alignment to standards, timely results, helpful reporting, helps inform instruction, etc.)*
6. What are some examples of district assessments that you have not found helpful in informing instructional practice, diagnosing student needs or predicting a student's later performance? How have they not been helpful? How could they be changed to be more helpful? *(Moderator prompt: assessment not aligned to standards, results not timely or in a helpful format, not designed to inform instruction, etc.)*
7. What district assessments, if any, would you suggest the district continue to administer as it does today? Why?
8. What district assessments, if any, would you suggest the district consider eliminating from the assessment program? Why?
9. What district assessments do you think need significant changes? Why? *(Moderator prompt: improve alignment or reporting, reduce frequency of administration, limit the grades or subject areas assessed, focus on a smaller subset of students who must take the assessment, etc.)*
10. Do you see any current gaps in the assessment program that the district should address?
11. What are some changes that occur in your classroom as testing windows approach? Are these positive or negative changes in your opinion?
12. What are some changes that occur in your classroom during the testing windows? Are these positive or negative changes?
13. Does the administration of state and/or districtwide tests take too much time, not enough time, or just enough time in your opinion? If too much, what are some factors that contribute to long testing days? *(Moderator prompt: lack of proper planning, technology failures, student discipline problems, etc.)*
14. Are there any other comments you have about district and state assessments?

APPENDIX B: CASE STUDIES

In this appendix, we provide a brief snapshot of each of the districts in our sample. The case studies provide contextual information as well as data highlighting the number of standardized test administrations and the test completion time for district- and state-mandated tests in each school district. Additionally, we present teacher perspectives captured through our focus groups. We have used pseudonyms to protect the anonymity of the districts participating in our study.

OVERTON: A Small, High-Poverty School District without 1:1 Technology

Overton serves a rural area with a population of approximately 8,000 people. The total student population in the district was roughly 1,500 during the 2014-2015 school year. A high percentage of Overton students live in poverty: over 95% of students in the district received free or reduced-price lunch.[29] The district is historically low performing. Since the 2008-2009 school year, the district has consistently earned either a "D" or an "F" rating or its equivalent. In 2015, the district was rated a "D" based on 2014-2015 data. Taking all federal, state, and local revenue into account, Overton received nearly $11,000 per pupil.

TESTING IN THE DISTRICT

Overton students in 2014-2015 took 256 state- and district-mandated assessments in K-12, counting all administrations of each test. Of the 256 tests, only 19% were state-mandated. The number of state and district assessments given in each grade averaged 20 but ranged from 0-31, with the greatest amount of state and district testing occurring in grades 5 and 8 (31 tests each). District testing accounted for 26 of the test administrations in these grades. Grade 3 students took the greatest number of state-mandated tests, with four state tests administered a total of eight times. District testing varied across the remaining grades, but every grade from K-8 took at least 21 district-mandated tests. High school students took no district-mandated tests.

On average, Overton students spent 15.6-16.2 hours taking state- and district-mandated tests in K-12 in 2014-2015. Students in grades 5 and 8 spent the greatest amount of time testing overall (28-30 hours and 28.83–30.83 hours, respectively). They also spent the most time taking district tests (both 15-17 hours). Grades 3, 6, and 7 spent about as much time on state tests as on district tests. Students in high school only took state tests. Students in grades K-2 and 4 spent more time on district tests than on state tests.

Case Study Table 1 shows the total number of tests and total completion time in each grade, broken out by district and state. Case Study Table 2 shows the number of tests that students took per grade in 2014-2015. For a test given multiple times a year, each administration is counted. This means that a total of four tests in a given grade may be four different tests or the same test administered four times a year. To clarify this, we have listed the name of the test with the number of administrations in parentheses in order to show how the total was derived.

[29] In 2014-2015, Overton began participating in the Community Eligibility Provision (CEP), which enables the district to provide free lunches to 100% of the district's students without collecting income data from all parents. Because districts can qualify for CEP for schools with a poverty rate as low as 40%, CEP makes precise poverty rates difficult to know. In 2013-2014, the last year before CEP went into effect in Mississippi, Overton had a free or reduced-price lunch percentage of over 95%.

APPENDIX B: CASE STUDIES, CONTINUED

OVERTON: A Small, High-Poverty School District without 1:1 Technology

CASE STUDY TABLE 1: Number of Tests and Completion Time by Grade

Grade	K	1	2	3	4	5	6	7	8	9	10	11	12
# of State Tests Given	5	3	3	8	4	5	4	4	5	3	2	2	0
Time on State Tests (hrs)	2.7	1	1	11.58	10	13	10.83	10.83	13.83	7.33-9.33	5.75	4.92	0
# of District Tests Given	21	21	21	21	24	26	24	24	26	0	0	0	0
Time on District Tests (hrs)	11	11	11	11	12	15-17	12	12	15-17	0	0	0	0
# of Total Tests Given	**26**	**24**	**24**	**29**	**28**	**31**	**28**	**28**	**31**	**3**	**2**	**2**	**0**
Total Time on Tests (hrs)	**13.7**	**12**	**12**	**22.58**	**22**	**28-30**	**22.83**	**22.83**	**28.83-30.83**	**7.33-9.33**	**5.75**	**4.92-6.92**	**0**

APPENDIX B: CASE STUDIES, CONTINUED

OVERTON: A Small, High-Poverty School District without 1:1 Technology

CASE STUDY TABLE 2: Test Administrations per Grade

Grade	Pre-K	K	1	2	3	4	5
State	N/A	1. MKAS[2] (2) 2. STAR Reading (3)	1. STAR Reading (3)	1. STAR Reading (3)	1. MKAS[2] (1) 2. STAR Reading (3) 3. PARCC ELA (2) 4. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)
District		3. STAR Reading (6) 4. STAR Math (9) 5. i-Ready ELA (3) 6. i-Ready Math (3)	2. STAR Reading (6) 3. STAR Math (9) 4. i-Ready ELA (3) 5. i-Ready Math (3)	2. STAR Reading (6) 3. STAR Math (9) 4. i-Ready ELA (3) 5. i-Ready Math (3)	5. STAR Reading (6) 6. STAR Math (9) 7. i-Ready ELA (3) 8. i-Ready Math (3)	3. STAR Reading (9) 4. STAR Math (9) 5. i-Ready ELA (3) 6. i-Ready Math (3)	4. STAR Reading (9) 5. STAR Math (9) 6. i-Ready ELA (3) 7. i-Ready Math (3) 8. Case21 Science (2)
TOTAL		**26**	**24**	**24**	**29**	**28**	**31**

Grade	6	7	8	9	10	11	12
State	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)	1. PARCC Algebra I (2) 2. Biology 1 (1)	1. PARCC English II (2)	1. US History (1) 2. ACT (1)	
District	3. STAR Reading (9) 4. STAR Math (9) 5. i-Ready ELA (3) 6. i-Ready Math (3)	3. STAR Reading (9) 4. STAR Math (9) 5. i-Ready ELA (3) 6. i-Ready Math (3)	4. STAR Reading (9) 5. STAR Math (9) 6. i-Ready ELA (3) 7. i-Ready Math (3) 8. Case21 Science (2)				
TOTAL	**28**	**28**	**31**	**3**	**2**	**2**	**0**

OTHER TESTS SOME, BUT NOT ALL, STUDENTS TAKE

For context purposes, we list below other tests that some students in Overton take in a school year.

- **Teacher-Created Tests.** Students at all grade levels routinely take tests created by their individual teachers to measure mastery of the curriculum. Students may also take tests created by groups of teachers, although no group tests were noted by Overton teachers in our focus groups.
- **SATP2 Re-tests.** Students who did not pass the 2013-2014 Algebra I, English II, Biology I, or U.S. History SATP2 exams were allowed to re-test two times.
- **Career Planning and Assessment System (CPAS).** The CPAS exam is given to vocational students in their senior year of high school. It is administered twice.

APPENDIX B: CASE STUDIES, CONTINUED

OVERTON: A Small, High-Poverty School District without 1:1 Technology

- **Mississippi Alternate Assessment of Extended Curriculum Frameworks (MAAECF).** The MAAECF was given to all students with disabilities who were unable to participate in the regular state assessment in 2014-2015.
- **National Assessment of Educational Progress (NAEP).** NAEP is administered to a random group of students in grades 4, 8, and 12 every few years.
- **World-Class Instructional Design Assessment (WIDA).** The WIDA was administered to all students entering the district for the first time in 2014-2015 who reported speaking another language at home. This test was administered to determine students' eligibility for the English language development program.
- **Advanced Placement (AP) Tests.** AP tests are administered to students in advanced placement courses to determine if students can receive college credit.
- **PSAT.** The PSAT is typically administered to tenth grade students. The test is used to identify National Merit Scholars.
- **ReadWorks.** ReadWorks is a free, online resource that enables teachers to find reading passages and comprehension questions aligned to the state standards. It is not a testing program, but teachers can use the passages and questions as part of teacher-created tests.

TEACHER PERSPECTIVES ON TESTING

I. Elementary Teachers

The perspectives of Overton elementary teachers about district assessments varied greatly. Some teachers expressed satisfaction with aspects of the district assessment program, while some expressed extreme dissatisfaction with the same aspects. For instance, one elementary teacher recommended continuing the STAR test because the teacher thought it was useful in identifying student skill deficiencies, while another elementary teacher recommended eliminating STAR tests because the teacher felt it did not accurately measure student readiness for the state ELA test. Furthermore, one teacher thought monthly STAR testing was beneficial while another found it detrimental. Finally, teachers were divided about whether the district tested too much or just enough. Two teachers thought that district testing was enough; one of these teachers indicated testing had been too much in the past but had been reduced at her school as of the interview. Only one of the teachers we interviewed thought that students experienced too much testing. This same, teacher, however, thought that no tests should be eliminated. These differences in perspectives may indicate a difference in how tests and data are used at various elementary schools across the district—even those tests administered districtwide—since each of the three teachers in the focus group worked at different schools.

As for state assessments, the elementary teachers were primarily concerned about a lack of clear communication about the 2015-2016 MAP assessment as of the date of the focus groups in November 2015. For two years in a row, none of the teachers knew what to expect from the state assessments. They viewed all other assessments as tools that should lead students to be successful on the state exams but found it hard to leverage other assessment data when they did not know what the state test would be like.

APPENDIX B: CASE STUDIES, CONTINUED

OVERTON: A Small, High-Poverty School District without 1:1 Technology

II. Middle School Teachers

Middle school teachers in Overton believed there was the "right amount" of testing at the middle school level, but they still had several concerns about both district and state testing. First, they felt that the district changed testing programs too often. One teacher commented that the district adopts a new progress-monitoring assessment each year, which is very disruptive to teachers. Teachers expressed the need for more professional development on the district assessments in place now so that they could utilize all the tools available within the testing software. Secondly, they believed there was too much ambiguity in the test questions in the district item bank that they must use for their nine-weeks assessments. They thought this led to unfair assessments. Additionally, they wanted access to detailed student performance data from the STAR assessments so that they could use it in making instructional decisions. Without the data, the STAR tests were of little value to them. Finally, they spoke strongly about needing data coaching to support their use of any assessment. Teachers wanted support not only on better utilizing testing software, but also on how to analyze the data that it provides. They also sought professional development on how to write better test questions aligned to state tests for their own teacher-created tests.

Overton middle school teachers held the same frustrations about state testing as teachers across the district. These frustrations include the rapid change in state tests and a lack of PARCC data, despite it being over six months after the test was administered at the time of the focus group. One middle school teacher was very vocal about wanting to entirely eliminate state testing because students constantly underperform. Middle school teachers also wanted more teacher input in state testing procedures at the school level to ensure that the process runs more smoothly.

III. High School Teachers

Because there are no high school district-mandated assessments, the high school educators (2 teachers, 1 administrator) we interviewed were mostly concerned about the state-mandated assessments. High school educators believed PARCC required too much testing. They also expressed a great deal of dissatisfaction with the lack of communication from MDE regarding state assessments. Changing assessments caused great confusion among teachers and hampered their ability to prepare their students. They expressed frustration at rumors that students would not need to pass the assessments to graduate. Students questioned whether they would have to take the test at all. Teachers feared being evaluated based on a test that students had no stake in.

APPENDIX B: CASE STUDIES, CONTINUED

HILLSIDE: A Small, Lower-Poverty School District with 1:1 Technology

Hillside serves a rural area with a population of approximately 15,000 people. The total student population in the district was roughly 2,700 during the 2014-2015 school year. Approximately 60% of Hillside students receive free or reduced-price lunch. The district is a high-performing district with an "A" accountability rating. Taking all federal, state, and local revenue into account, Hillside received nearly $8,700 per pupil.

TESTING IN THE DISTRICT

Hillside students in 2014-2015 took 144 state- and district-mandated assessments in K-12, counting all administrations of each test. Of the 144 tests, 33% were state-mandated. The number of state and district assessments given in each grade averaged 11 but ranged from 0-23, with the greatest amount of state and district testing occurring in grades 2 (18) and 3 (23). District testing accounted for 15 of the test administrations in these grades. In addition to taking the greatest number of tests overall, grade 3 students also took the greatest number of state-mandated tests, with four state tests administered a total of eight times. District testing varied across the remaining grades, but every grade from K-8 took at least six district-mandated tests. District testing at the high school level was confined to grades 9-10, but Hillside eleventh and twelfth graders are likely to take tests designed for a particular group of students, such as exams for an advanced diploma.

On average, Hillside students spent a total of 13 hours taking state- and district-mandated tests in K-12 in 2014-2015. Students in grades 8 and 10 spent the greatest amount of time testing overall and on both state and district tests. Students in grades K-2 and 9 spent more time on district than state tests; all other students spent more time on state tests.

Case Study Table 3 shows the total number of tests and total completion time in each grade, broken out by district and state. Case Study Table 4 shows the number of tests that students took per grade in 2014-2015. For a test given multiple times a year, each administration is counted. This means that a total of four tests in a given grade may be four different tests or the same test administered four times a year. To clarify this, we have listed the name of the test with the number of administrations in parentheses in order to show how the total was derived.

APPENDIX B: CASE STUDIES, CONTINUED

HILLSIDE: A Small, Lower-Poverty School District with 1:1 Technology

CASE STUDY TABLE 3: Number of Tests and Completion Time by Grade

Grade	K	1	2	3	4	5	6	7	8	9	10	11	12
# of State Tests Given	5	3	3	8	4	5	4	4	5	1	5	1	0
Time on State Tests (hrs)	2.7	1	1	11.58	10	13	10.83	10.83	13.83	2-4	13.08-15.08	2.92	0
# of District Tests Given	12	9	15	15	10	6	7	6	9	3	4	0	0
Time on District Tests (hrs)	3.25	3	6	6	7	6	7	6	11.83	8	11.75	0	0
# of Total Tests Given	**17**	**12**	**18**	**23**	**14**	**11**	**11**	**10**	**14**	**4**	**9**	**1**	**0**
Total Time on Tests (hrs)	**5.95**	**4**	**7**	**17.58**	**17**	**19**	**17.83**	**16.83**	**25.67**	**10-12**	**24.83-26.83**	**2.92**	**0**

APPENDIX B: CASE STUDIES, CONTINUED

HILLSIDE: A Small, Lower-Poverty School District with 1:1 Technology

CASE STUDY TABLE 4: Test Administrations per Grade

Grade	Pre-K	K	1	2	3	4	5
State	N/A	1. MKAS[2] (2) 2. STAR Reading (3)	1. STAR Reading (3)	1. STAR Reading (3)	1. MKAS[2] (1) 2. STAR Reading (3) 3. PARCC ELA (2) 4. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)
District		3. DIBELS (3) 4. STAR Math (6) 5. STAR Reading (3)	2. STAR Math (6) 3. STAR Reading (3)	2. STAR Math (6) 3. STAR Reading (3) 4. SRI (6)	5. STAR Math (6) 6. STAR Reading (3) 7. SRI (6)	3. SRI (6) 4. NWEA MAP ELA (2) 5. NWEA MAP Math (2)	4. NWEA MAP ELA (3) 5. NWEA MAP Math (3)
TOTAL		**17**	**12**	**18**	**23**	**14**	**11**

Grade	6	7	8	9	10	11	12
State	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)	1. Biology I (1)	1. PARCC English II (2) 2. PARCC Algebra I (2) 3. US History (1)	1. ACT (1)	
District	3. NWEA MAP ELA (3) 4. NWEA MAP Math (3) 5. Orleans-Hanna (1)	3. NWEA MAP ELA (3) 4. NWEA MAP Math (3)	4. NWEA MAP ELA (3) 5. NWEA MAP Math (3) 6. Curriculum ELA (1) 7. Curriculum Math (1) 8. Curr. Science (1)	2. Curriculum ELA (1) 3. Curriculum Math (1) 4. Curr. Biology (1)	4. Curriculum ELA (1) 5. Curr. Math II (1) 6. Curr. Chemistry (1) 7. Curr. World Hist.(1)		
TOTAL	**11**	**10**	**14**	**4**	**9**	**1**	**0**

OTHER TESTS SOME, BUT NOT ALL, STUDENTS TAKE

For context purposes, we list below other tests that some students in Hillside take in a school year.

- **Teacher-Created Tests.** Students at all grade levels routinely take tests created by their individual teachers to measure mastery of the curriculum. Students may also take tests created by groups of teachers. For Hillside's specialized curriculum in grades 9-12 (referred to herein as "Curriculum"), teachers must submit student coursework as part of each Curriculum exam in order for students to receive an overall score in each Curriculum course. These coursework assignments, though part of the Curriculum exams, are developed and assigned by each teacher using guidelines provided by the Curriculum.

APPENDIX B: CASE STUDIES, CONTINUED

HILLSIDE: A Small, Lower-Poverty School District with 1:1 Technology

- **Advanced Curriculum Assessments.** For students to receive the Advanced Curriculum diploma, which is an honors diploma in the Curriculum program, they must take Advanced Curriculum assessments, which are for advanced courses beyond the standard Curriculum course of study. Hillside students in eleventh and twelfth grades may choose to pursue the Advanced Curriculum diploma.
- **SATP2 Re-tests.** Students who did not pass the 2013-2014 Algebra I, English II, Biology I, or U.S. History SATP2 exams were allowed to re-test two times.
- **Edsphere.** Edsphere is a computerized personalized-learning program for ELA that contains diagnostic assessments with aligned skill-building lessons for remediation. This program, and its assessment, is used at the discretion of Hillside teachers.
- **Advanced Placement (AP) Tests.** AP tests are administered to students in advanced placement courses to determine if students can receive college credit. AP Calculus is offered at Hillside High School.
- **Mississippi Alternate Assessment of Extended Curriculum Frameworks (MAAECF).** The MAAECF was given to all students with disabilities who were unable to participate in the regular state assessment in 2014-2015.
- **National Assessment of Educational Progress (NAEP).** NAEP is administered to a random group of students in grades 4, 8, and 12 every few years.
- **World-Class Instructional Design Assessment (WIDA).** The WIDA was administered to all students entering the district for the first time in 2014-2015 who reported speaking another language at home. This test was administered to determine students' eligibility for the English language development program.
- **PSAT.** The PSAT is typically administered to tenth grade students. The test is used to identify National Merit Scholars.

TEACHER PERSPECTIVES ON TESTING

I. Elementary Teachers

All teachers at the elementary level thought students were overexposed to testing, despite the fact that Hillside had a low number of tests relative to others in our sample. Recent changes in district curricula, state standards, and state assessments were a source of extreme anxiety to teachers. They described an environment in which district and state testing had become so frequent that parents and students could not differentiate the purposes and importance of each of the tests.

Elementary teachers thought that the STAR Reading, NWEA MAP, and SRI assessments were valuable to their work because they provided immediate feedback, automatically grouped students for remediation, and easily tracked student growth. They also found STAR Reading to be

APPENDIX B: CASE STUDIES, CONTINUED

HILLSIDE: A Small, Lower-Poverty School District with 1:1 Technology

an accurate predictor of success on the third grade MKAS[2] test. Despite strongly liking each of these products individually, they felt that all of the products used together amounted to too much time taken away from instruction.

Elementary teachers additionally expressed great dissatisfaction with state-mandated testing. They had not yet received scores from the PARCC exam as of November 2015, and they did not know how to use the MKAS[2] cut score to make meaningful analyses about the performance of their students.

II. Middle School Teachers

Like their elementary counterparts, middle school teachers thought students were overexposed to testing. All expressed dissatisfaction with the state- and district-mandated assessments given to middle school students, although they more strongly disliked the state assessments. Middle school teachers believed that testing was necessary but felt that more emphasis on data analysis and use was required to make testing truly worthwhile. They felt that too many tests not only took away from instruction but also prevented them from making good use of any one test.

III. High School Teachers

Of all the Hillside teacher groups, high school teachers were the most satisfied with the current testing regimen in Hillside, even though they still had concerns about testing. Two of three high school teachers thought students were overexposed to testing, but the picture is more complicated at high school where "over-testing" can be ascribed to the double-testing for both Hillside's Curriculum tests and the state tests, rather than frequent progress-monitoring assessments.

An encouraging sign was that all high school teachers were satisfied with the Curriculum assessments. They thought these tests were helpful to their instruction and that student success on the assessments translated to success in college. Like elementary and middle school teachers, they were dissatisfied with the state assessments. They thought the time invested in preparing for and taking state assessments was wasted.

APPENDIX B: CASE STUDIES, CONTINUED

MANNEQUIN: A Mid-Sized, Lower-Poverty District without 1:1 Technology

Mannequin serves a community with a population of over 30,000 people. The total student population in the district was roughly 5,800 during the 2014-2015 school year. Approximately 40% of Mannequin students received free or reduced-price lunch. The district is a high-performing district with an "A" accountability rating. Taking all federal, state, and local revenue into account, Mannequin received nearly $8,500 per pupil.

TESTING IN THE DISTRICT

Mannequin students in 2014-2015 took 130 state- and district-mandated assessments in K-12, counting all administrations of each test. Of the 130 tests, 37% were state-mandated. The number of state and district assessments given in each grade averaged 10 but ranged from 0-15, with the greatest amount of district and state testing occurring in grades 3-8. Grade 3-8 students took 14-15 state- and district-mandated tests each year. High school students took far fewer district and state tests, but Mannequin high school students are more likely to take tests that are designed for a particular group of students, such as exams for advanced placement or career and technical education courses. Grade 3 students took the greatest number of state-mandated tests, with four state tests administered a total of eight times. District testing was consistent across each of the following grade spans: K-3 (three assessments with seven total administrations), 4-8 (four assessments with a total of ten administrations), and 9-12 (three assessments with a total of four administrations). The number of district-mandated math tests was greater than district-mandated ELA tests only because the district's K-3 ELA screener is considered a state test. The district did not mandate any assessments in other subject areas.

On average, Mannequin students spent 11.9-13.9 hours taking state- and district-mandated tests in 2014-2015. Fifth and eighth grade students spent the greatest amount of time testing. In grades 3, 5, 8, and 9-11,[30] students spent more time on state tests. Students in grade 4 spent comparable amounts of time on state and district tests. Students in grades K-2 and 6-7 spent more time on district tests.

Case Study Table 5 shows the total number of tests and total completion time in each grade, broken out by district and state. Case Study Table 6 shows the number of tests that students took per grade in 2014-2015. For a test given multiple times a year, each administration is counted. This means that a total of four tests in a given grade may be four different tests or the same test administered four times a year. To clarify this, we have listed the name of the test with the number of administrations in parentheses in order to show how the total was derived.

[30] This assumes ninth graders took both the PARCC Algebra I exam and the Biology I exam.

APPENDIX B: CASE STUDIES, CONTINUED

MANNEQUIN: A Mid-Sized, Lower-Poverty District without 1:1 Technology

CASE STUDY TABLE 5: Number of Tests and Completion Time by Grade

Grade	K	1	2	3	4	5	6	7	8	9	10	11	12
# of State Tests Given	5	3	3	8	4	5	4	4	5	3	2	2	0
Time on State Tests (hrs)	2.7	1	1	11.58	10	13	10.83	10.83	13.83	7.33-9.33	5.75	4.92	0
# of District Tests Given	7	7	7	7	10	10	10	10	10	3	1	0	0
Time on District Tests (hrs)	3.67-6	3.67-6	3.67-6	3.67-6	8.67-11	8.67-11	8.67-11	8.67-11	8.67-11	2.5-3.75	.833-1.25	0	0
# of Total Tests Given	**12**	**10**	**10**	**15**	**14**	**15**	**14**	**14**	**15**	**6**	**3**	**2**	**0**
Total Time on Tests (hrs)	**6.37-8.7**	**4.67-7**	**4.67-7**	**15.25-17.58**	**18.67-21**	**21.67-24**	**19.5-21.83**	**19.50-21.83**	**22.50-24.83**	**9.83-13.08**	**6.58-7**	**4.92-6.92**	**0**

APPENDIX B: CASE STUDIES, CONTINUED

MANNEQUIN: A Mid-Sized, Lower-Poverty District without 1:1 Technology

CASE STUDY TABLE 6: Test Administrations per Grade

Grade	Pre-K	K	1	2	3	4	5
State	N/A	1. MKAS[2] (2) 2. STAR Reading (3)	1. STAR Reading (3)	1. STAR Reading (3)	1. MKAS[2] (1) 2. STAR Reading (3) 3. PARCC ELA (2) 4. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)
District		1. STAR Math (3) 2. District ELA (2) 3. District Math (2)	2. STAR Math (3) 3. District ELA (2) 4. District Math (2)	2. STAR Math (3) 3. District ELA (2) 4. District Math (2)	5. STAR Math (3) 6. District ELA (2) 7. District Math (2)	3. i-Ready ELA (3) 4. i-Ready Math (3) 5. District ELA (2) 6. District Math (2)	4. i-Ready ELA (3) 5. i-Ready Math (3) 6. District ELA (2) 7. District Math (2)
TOTAL		**12**	**10**	**10**	**15**	**14**	**15**

Grade	6	7	8	9	10	11	12
State	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)	1. PARCC Algebra I (2) 2. Biology 1 (1)	1. PARCC English II (2)	1. US History (1) 2. ACT (1)	
District	3. i-Ready ELA (3) 4. i-Ready Math (3) 5. District ELA (2) 6. District Math (2)	3. i-Ready ELA (3) 4. i-Ready Math (3) 5. District ELA (2) 6. District Math (2)	4. i-Ready ELA (3) 5. i-Ready Math (3) 6. District ELA (2) 7. District Math (2)	3. Dist. English I (1) 4. Dist. Algebra I (2)	2. Dist. English II (1)		
TOTAL	**14**	**14**	**15**	**6**	**3**	**2**	**0**

OTHER TESTS SOME, BUT NOT ALL, STUDENTS TAKE

For context purposes, we list below other tests that some students take in a school year.

- **Teacher-Created Tests.** Students at all grade levels routinely take tests created by their individual teachers to measure mastery of the curriculum. Students may also take tests created by groups of teachers. Elementary school teachers in Mannequin work in grade-level teams to create weekly tests for each grade level. High school teachers in each subject area give quarterly assessments in addition to any tests mandated by the district.
- **SATP2 Re-tests.** Students who did not pass the 2013-2014 Algebra I, English II, Biology I, or U.S. History SATP2 exams were allowed to re-test two times.
- **Career Planning and Assessment System (CPAS).** The CPAS exam is given to vocational students in their senior year of high school. It is administered twice.

APPENDIX B: CASE STUDIES, CONTINUED

MANNEQUIN: A Mid-Sized, Lower-Poverty District without 1:1 Technology

- **Mississippi Alternate Assessment of Extended Curriculum Frameworks (MAAECF).** The MAAECF was given to all students with disabilities who were unable to participate in the regular state assessment in 2014-2015.
- **National Assessment of Educational Progress (NAEP).** NAEP is administered to a random group of students in grades 4, 8, and 12 every few years.
- **World-Class Instructional Design Assessment (WIDA).** The WIDA was administered to all students entering the district for the first time in 2014-2015 who reported speaking another language at home. This test was administered to determine students' eligibility for the English language development program.
- **Advanced Placement (AP) Tests.** AP tests are administered to students in AP courses to determine if students can receive college credit.
- **International Baccalaureate (IB) Tests.** IB tests are administered to students in IB courses to determine if students can receive college credit.
- **PSAT.** The PSAT is typically administered to tenth grade students. The test is used to identify National Merit Scholars.

TEACHER PERSPECTIVES ON TESTING

I. Elementary Teachers

Elementary teachers in Mannequin thought that students are given too many tests and that testing takes too much time in the school year. Their concerns were based on the district assessments and the third grade MKAS^2. They felt the district assessments were not aligned to their pacing guides and, as a result, to what was taught in the classroom. This makes data from the test of little instructional value to them. Teachers believed that data from STAR, i-Ready, and teacher-created weekly tests were sufficient for making instructional decisions in their classrooms. They valued the STAR and i-Ready exams highly because the results are immediate, allowing them to analyze data quickly and serve their students more effectively. Interestingly, they felt that a high-stakes test like MKAS^2 should be given in an earlier grade when students are still being taught to read and appropriate interventions can be provided.

II. Middle School Teachers

Mannequin's middle school teachers thought that students took too many state- and district-mandated tests even though Mannequin tests the least of all the districts studied. They reported the district-mandated i-Ready tests made many of the resources (e.g., computers, labs, and the media center) that would enhance their lessons unavailable due to the frequency of computerized testing. They have concerns about whether the i-Ready test data is valuable because they do not think that students take them seriously (they noted that this was the first year of i-Ready and that, as of October when the focus group occurred, it was too early to determine whether the cost in terms of time and resources would be worthwhile). One teacher estimated a test was given in the building somewhere on 36 days of the school year.

APPENDIX B: CASE STUDIES, CONTINUED

MANNEQUIN: A Mid-Sized, Lower-Poverty District without 1:1 Technology

The middle school teachers thought that the assessments they created in their own classrooms were most valuable to their practice because they provide instant feedback about student learning, and they are aligned to what has been taught in the classroom. They also thought that students take teacher-created assessments more seriously because they are more aligned to classroom lessons and students get to show what they have been taught.

III. High School Teachers

High school teachers thought that students took too many state- and district-mandated tests. Teachers have to cover classes for others during test administration days, and it takes away valuable instruction time. Additionally, they thought that state tests, particularly the Mississippi-created subject-area tests, set a low bar for students. Because all other tests are created to ensure that students can be successful on the state tests, other tests reinforce the low bar. They think that students are not being challenged, and teachers are not able to teach more rigorous skills and content because the state tests dictate the focus of their classrooms.

The high school teachers thought that their own assessments were most valuable because they were rigorous and more reflective of what their students were capable of doing. They thought that district-created assessments are more helpful when student data from the exams is compiled, analyzed, and used in a collaborative meeting where teachers discussed the implications. Teachers with less experience also found district-created exams more helpful because they did not have as many teacher-created tests of their own to rely upon.

APPENDIX B: CASE STUDIES, CONTINUED

SUNSET: A Mid-Sized, High-Poverty District with Partial 1:1 Technology

The Sunset School District serves a community with a population of over 30,000 people. The total student population in the district was roughly 5,600 during the 2014-2015 school year. A high percentage of Sunset students live in poverty: over 90% of students qualified for free or reduced-price lunch.[31] The district is historically low-performing. Since the 2008-2009 school year, the district has consistently held either a "D" or an "F" rating or its equivalent. In 2015, the district was rated an "F" based on 2014-2015 data. Taking all federal, state, and local revenue into account, Sunset received nearly $8,900 per pupil.

TESTING IN THE DISTRICT

Sunset students in 2014-2015 took 269 state- and district-mandated assessments in K-12, counting all administrations of each test. Of the 269 tests, only 18% were state-mandated. The number of state and district assessments given in each grade averaged 21 but ranged from 6-27, with the greatest amount of state and district testing occurring in grades 3 and 5 (27 tests each). District testing accounted for 19 and 22, respectively, of the test administrations in these grades. In addition to tying for the greatest number of tests overall, grade 3 students also took the greatest number of state-mandated tests, with four state tests administered a total of eight times. District testing varied across the remaining grades, but every grade from 2-12 took at least 19 district-mandated tests.

On average, Sunset students spent 48.9-49.2 hours taking state- and district-mandated tests in K-12 in 2014-2015. Students in grades 5 and 8 spent the greatest amount of time testing overall (63 and 62.8 hours, respectively) and on state tests (13 and 13.8 hours, respectively); students in grade 5 are also tied for the most time spent on district tests (50 hours; tied with grade 4). All grades but kindergarten and grade 1 spent more time on district tests (49 hours annually) than on state tests (average time between 7.1-7.4 hours annually).

Case Study Table 7 shows the total number of tests and total completion time in each grade, broken out by district and state. Case Study Table 8 shows the number of tests that students took per grade in 2014-2015. For a test given multiple times a year, each administration is counted. This means that a total of four tests in a given grade may be four different tests or the same test administered four times a year. To clarify this, we have listed the name of the test with the number of administrations in parentheses in order to show how the total was derived.

[31] In 2014-2015, Sunset began participating in the Community Eligibility Provision (CEP), which enables the district to provide free lunches to 100% of the district's students without collecting income data from all parents. Because districts can qualify for CEP for schools with a poverty rate as low as 40%, CEP makes precise poverty rates difficult to know. In 2013-2014, the last year before CEP went into effect in Mississippi, Sunset had a free or reduced-price lunch percentage of approximately 90%.

APPENDIX B: CASE STUDIES, CONTINUED

SUNSET: A Mid-Sized, High-Poverty District with Partial 1:1 Technology

CASE STUDY TABLE 7: Total Number of Tests and Total Completion Time by Grade

Grade	K	1	2	3	4	5	6	7	8	9	10	11	12
# of State Tests Given	5	3	3	8	4	5	4	4	5	3	2	2	0
Time on State Tests (hrs)	2.7	1	1	11.58	10	13	10.83	10.83	13.83	7.33-9.33	5.75	4.92	0
# of District Tests Given	3	3	19	19	22	22	19	19	19	19	19	19	19
Time on District Tests (hrs)	1	1	49	49	50	50	49	49	49	49	49	49	49
# of Total Tests Given	**8**	**6**	**22**	**27**	**26**	**27**	**23**	**23**	**24**	**22**	**21**	**21**	**19**
Total Time on Tests (hrs)	**3.7**	**2**	**50**	**60.58**	**60**	**63**	**59.83**	**59.83**	**62.83**	**56.33-58.33**	**54.75**	**53.92-55.92**	**49**

APPENDIX B: CASE STUDIES, CONTINUED

CASE STUDY TABLE 8: Test Administrations per Grade

Grade	Pre-K	K	1	2	3	4	5
State	N/A	1. MKAS^2 (2) 2. STAR Reading (3)	1. STAR Reading (3)	1. STAR Reading (3)	1. MKAS^2 (1) 2. STAR Reading (3) 3. PARCC ELA (2) 4. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)
District		3. STAR Math (3)	2. STAR Math (3)	2. STAR Math (3) 3. District ELA (4) 4. District Math (4) 5. District Science (4) 6. District History (4)	5. STAR Math (3) 6. District ELA (4) 7. District Math (4) 8. District Science (4) 9. District History (4)	3. STAR Math (3) 4. STAR Reading (3) 5. District ELA (4) 6. District Math (4) 7. District Science (4) 8. District History (4)	4. STAR Math (3) 5. STAR Reading (3) 6. District ELA (4) 7. District Math (4) 8. District Science (4) 9. District History (4)
TOTAL		**8**	**6**	**22**	**27**	**26**	**27**

Grade	6	7	8	9	10	11	12
State	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2)	1. PARCC ELA (2) 2. PARCC Math (2) 3. MST (1)	1. PARCC Algebra I (2) 2. Biology I (1)	1. PARCC English II (2)	1. US History (1) 2. ACT (1)	
District	3. STAR Reading (3) 4. District ELA (4) 5. District Math (4) 6. District Science (4) 7. District History (4)	3. STAR Reading (3) 4. District ELA (4) 5. District Math (4) 6. District Science (4) 7. District History (4)	4. STAR Reading (3) 5. District ELA (4) 6. District Math (4) 7. District Science (4) 8. District History (4)	3. STAR Reading (3) 4. District ELA (4) 5. District Math (4) 6. District Science (4) 7. District History (4)	3. STAR Reading (3) 4. District ELA (4) 5. District Math (4) 6. District Science (4) 7. District History (4)	2. STAR Reading (3) 3. District ELA (4) 4. District Math (4) 5. District Science (4) 6. District History (4)	1. STAR Reading (3) 2. District ELA (4) 3. District Math (4) 4. District Science (4) 5. District History (4)
TOTAL	**23**	**23**	**24**	**22**	**21**	**21**	**19**

OTHER TESTS SOME, BUT NOT ALL, STUDENTS TAKE

For context purposes, we list below tests that some students take in a school year.

- **Teacher-Created Tests.** Students at all grade levels routinely take tests created by their individual teachers to measure mastery of the curriculum. Students may also take tests created by groups of teachers, although no group tests were noted by Sunset teachers in our focus groups.
- **A+ Learning.** A+ Learning is a computerized personalized-learning program for all subjects that creates individualized lessons for students and has an assessment component. Middle school students enrolled in computer lab class use this program.
- **USA Test Prep.** USA Test Prep is a computerized testing program that individual teachers may use to create tests or generate bell-ringer questions. Sunset High School teachers use this program to create "checkpoint" assessments throughout the school year.
- **SATP2 Re-tests.** Students who did not pass the 2013-2014 Algebra I, English II, Biology I, or U.S. History SATP2 exams were allowed to re-test two times.

APPENDIX B: CASE STUDIES, CONTINUED

SUNSET: A Mid-Sized, High-Poverty District with Partial 1:1 Technology

- **Advanced Placement (AP) Tests.** AP tests are administered to students in advanced placement courses to determine if students can receive college credit. Several AP courses are offered at Sunset High School in each of the four core subject areas.
- **Mississippi Alternate Assessment of Extended Curriculum Frameworks (MAAECF).** The MAAECF was given to all students with disabilities who were unable to participate in the regular state assessment in 2014-2015.
- **National Assessment of Educational Progress (NAEP).** NAEP is administered to a random group of students statewide in grades 4, 8, and 12 every few years.
- **World-Class Instructional Design Assessment (WIDA).** The WIDA was administered to all students entering the district for the first time in 2014-2015 who reported speaking another language at home. This test was administered to determine students' eligibility for the English language development program.
- **PSAT.** The PSAT is typically administered to tenth grade students. The test is used to identify National Merit Scholars.

TEACHER PERSPECTIVES ON TESTING

I. Elementary Teachers

Sunset elementary teachers had many different concerns about testing, although they had mixed opinions on whether students were "over-tested." First, they noticed the increase in testing at the elementary level over time. This increase is particularly hard on grades 3-5, and they worry that third graders feel "burnt out," especially due to the high-stakes MKAS[2]. One of their biggest concerns had to do with district testing and its alignment to district curriculum maps. They felt that there were alignment issues between the curriculum maps and the assessments in each subject. Additionally, K-1 teachers expressed reservations about plans to extend district assessments to K-1. Teachers worried that the tests would not be written to an appropriate reading level for students still learning basic literacy skills. Furthermore, they expressed pressure to finish teaching the year's objectives in the first three quarters of the year in order to spend the last nine weeks doing review and test prep. Finally, elementary teachers expressed dissatisfaction with state-mandated testing, saying that they never received scores from the PARCC exam after feeling that the test itself was too long. They felt unprepared for PARCC because they lacked resources about what the test would be like. They were experiencing similar feelings about MAP when the focus groups occurred in November 2015.

Elementary teachers did have some positive things to say about testing. They acknowledged testing's importance to instructional decision-making and felt that their colleagues thought the same. They found district testing more helpful than harmful. Elementary teachers were very positive that the STAR Reading and Math assessments were valuable to their work because they provide immediate feedback, automatically group students for remediation, and easily track student growth. They also found STAR Reading to be an accurate predictor of success on the third grade MKAS[2].

APPENDIX B: CASE STUDIES, CONTINUED

SUNSET: A Mid-Sized, High-Poverty District with Partial 1:1 Technology

II. Middle School Teachers

Of all of the Sunset teachers interviewed for this report, middle school teachers were the most positive about testing. They did not voice strong complaints about the prevalence of either state- or district-mandated testing. The largest area of concern was the alignment of the district's nine-weeks assessments to curriculum maps. Like teachers at the elementary level, middle school teachers reported problems with the nine-weeks assessments covering content that was not on the curriculum map until later in the year. This mismatch caused feelings of anxiety among students and feelings of unfairness among teachers. They wanted more communication from the instructional strategists about what objectives would be covered by the exams so that they could better prepare students. One teacher also expressed concern that the questions on the district assessments relied too heavily on "regurgitation" of information rather than critical thinking.

Like the elementary teachers, Sunset middle school teachers reported that they must teach all of their objectives in the first three quarters of the year so that they could spend the entire last nine weeks reviewing. Teachers at one middle school stated that some of the core subject teachers even rotated to new students during this time to give students a different "perspective."

III. High School Teachers

High school teachers interviewed were mostly positive about testing but did express two important concerns. First, high school teachers felt that the STAR Reading exam had not been properly implemented at the high school. They had serious questions about its accuracy because they observe many students not taking the test seriously and simply marking answers to complete it. Furthermore, teachers reported the test was administered too late in the semester (October) and that not all teachers received the data, making it not helpful to them as a tool. Second, teachers reported that the district benchmark assessments are given too frequently at the high school, which is on a 4X4 block schedule. To simulate nine-weeks assessments, the district benchmarks are given every four and a half weeks in state-tested subjects. Between these tests, teachers are required to give "checkpoint" assessments. Considering both the district assessments and the checkpoint assessments, these teachers are required to test every two weeks. Teachers felt that this constant testing takes away from instructional time, which is already limited due to block scheduling. They felt the district should move to a pre-assessment, midterm assessment, and end-of-course exam. Lastly, state-tested teachers reported that they must teach all of their objectives in the first three quarters of the semester because they spend the last quarter reviewing for state tests. With the frequent assessments and pressure to finish teaching content early, teachers feel a lot of pressure to cover every objective.

APPENDIX C: MASTER TABLE OF ASSESSMENTS

Below is a master list of all assessments included in this report with administration frequencies and times used in our analyses. Assessments and administration frequencies were self-reported by each district on the survey tool in Appendix A. Times for state assessments were confirmed with Mississippi Department of Education officials.[32] As a result, we may have used a different time than that reported by the district. These times are as follows:

- For timed state tests, we have used official maximum time allotted; these tests include the MKAS[2]; PARCC; the Mississippi Science Test, Second Edition; and the ACT.
- For the two untimed state tests—the Biology I test and the U.S. History test—we used a state-provided average range of time.

For district tests created by the district, we use times self-reported by the district.

For district tests purchased from a vendor, we use times reported by the vendor in order to keep times consistent across districts that use the same products. Vendor test times used are as follows:

- For STAR Reading and STAR Math, we use an expected test time of 20 minutes per exam, as recommended by Renaissance Learning.[33]
- For i-Ready Reading and Math, we use an average test time of 60 minutes. Curriculum Associates reported that schools should allot two 45-minute class periods for the administration of each exam but that "many students may complete it in the first period."[34] We chose 60 minutes because it is the midway point between 45 minutes and 90 minutes and because it fell within the district-reported range for this test in multiple districts.
- For DIBELS, we use an expected test time of 5 minutes, as estimated by the University of Oregon and Hillside administrators.[35]
- For SRI, we use 30 minutes, as recommended by Houghton Mifflin Harcourt.[36]
- For NWEA MAP ELA and Math, we use an expected test time of 60 minutes per exam, as recommended by NWEA.[37]
- For the Orleans-Hanna Algebra Prognosis Test, we use an expected test time of 60 minutes, as recommended by Pearson.[38]
- For Case21, we use times reported by the district.

Hillside administered standardized assessments created by their curriculum vendor, which we refer to simply as "Curriculum" to maintain district anonymity. For these exams, we use information provided by the district as follows:

- For the eighth grade exams, we use administration times provided by Hillside Middle School for Curriculum English, math, and science. There are two tests for each of these subjects within one testing window.
- For the Curriculum exams administered in ninth and tenth grades, we use administration times provided by Hillside High School for each course in the standard sequence for Hillside High School students. The assessment for each course, except for Math II, includes one or two written tests given within the Curriculum testing window as well as the submission of graded coursework. We do not include time related to graded coursework that is submitted to the Curriculum vendor as this work is part of the normal operations of the course.

Sunset administered district-mandated benchmark tests in all four of the core subject areas in 2014-2015. These tests were hybrids in that they were created by the district's instructional strategists using an assessment creation tool and items owned by Educational Leadership Solutions (ELS). Because district staff created these tests to their specifications, we use the district-reported test time of 180 minutes per ELS exam.

Although some school districts in our sample had pre-K programs, we do not include pre-K test data in any of our tables or analyses because pre-K is not mandatory in Mississippi and not all districts have programs.

[32] Vincent Segalini (Mississippi Department of Education), e-mail message to R. Canter, June 16, 2016.
[33] Michael Bischoff (Renaissance Learning), online chat with R. Canter, June 16, 2016.
[34] Susan McCormack (Curriculum Associates), e-mail message to R. Canter, June 16, 2016.
[35] University of Oregon. 2016. "What are DIBELS?" Accessed December 22, 2017. https://dibels.uoregon.edu/assessment/dibels.
[36] Houghton Mifflin Harcourt. 2016. "About HMH Reading Inventory." Accessed December 21, 2017. http://www.hmhco.com/products/assessment-solutions/literacy/sri-index.htm.
[37] Northwest Evaluation Association (NWEA). 2015. *How MAP Enables Teachers and School Leaders to Make a Difference*. Brochure, Portland, OR: NWEA.
[38] Pearson. 2016. "Orleans-Hanna Algebra Prognosis Test, Third Edition." Accessed December 21, 2017. http://www.pearsonassessments.com/learningassessments/products/100000448/orleans-hanna-algebra-prognosis-test-third-edition.html.

MASTER TABLE OF ASSESSMENTS, CONTINUED

Grade Level	Assessment	District or State Mandated?	Paper/Computerized	Test Frequency	Time for 1 Test (minutes)	Time for All Administrations
K	MKAS[2]	State	Computerized	2	51	102
K	STAR Reading	State[39]	Computerized	3	20	60
K	STAR Reading	Overton	Computerized	6	20	120
K	STAR Reading	Hillside	Computerized	3	20	60
K	STAR Math	Overton	Computerized	9	20	180
K	STAR Math	Hillside	Computerized	3	20	60
K	STAR Math	Mannequin	Computerized	3	20	60
K	STAR Math	Sunset	Computerized	3	20	60
K	i-Ready ELA	Overton	Computerized	3	60	180
K	i-Ready Math	Overton	Computerized	3	60	180
K	DIBELS	Hillside	Paper	3	5	15
K	ELA Common Assessment	Mannequin	Paper	2	40-75	80-150
K	Math Common Assessment	Mannequin	Paper	2	40-75	80-150
1	STAR Reading	State	Computerized	3	20	60
1	STAR Reading	Overton	Computerized	6	20	120
1	STAR Reading	Hillside	Computerized	3	20	60
1	STAR Math	Overton	Computerized	9	20	180
1	STAR Math	Hillside	Computerized	3	20	60
1	STAR Math	Mannequin	Computerized	3	20	60
1	STAR Math	Sunset	Computerized	3	20	60
1	i-Ready ELA	Overton	Computerized	3	60	180
1	i-Ready Math	Overton	Computerized	3	60	180
1	ELA Common Assessment	District	Paper	2	40-75	80-150
1	Math Common Assessment	District	Paper	2	40-75	80-150
2	STAR Reading	State	Computerized	3	20	60
2	STAR Reading	Overton	Computerized	6	20	120
2	STAR Reading	Hillside	Computerized	3	20	60
2	STAR Math	Overton	Computerized	9	20	180
2	STAR Math	Hillside	Computerized	3	20	60
2	STAR Math	Mannequin	Computerized	3	20	60
2	STAR Math	Sunset	Computerized	3	20	60

[39] Every district in our sample used STAR Reading as its state-mandated screener for the *Literacy-Based Promotion Act* (LBPA). The LBPA requires students to be screened for reading ability three times a year—beginning, middle, and end. Schools in 2014-2015 were also required to use the MKAS[2] for the K-entry and grade 3 exit screener. As a result, we list STAR Reading as a state test for three administrations per year in K-3. If districts mandated additional administrations, these administrations are listed separately.

MASTER TABLE OF ASSESSMENTS, CONTINUED

Grade Level	Assessment	District or State Mandated?	Paper/Computerized	Test Frequency	Time for 1 Test (minutes)	Time for All Administrations
2	i-Ready ELA	Overton	Computerized	3	60	180
2	i-Ready Math	Overton	Computerized	3	60	180
2	SRI	Hillside	Computerized	6	30	180
2	ELA Common Assessment	Mannequin	Paper	2	40-75	80-150
2	Math Common Assessment	Mannequin	Paper	2	40-75	80-150
2	ELS - ELA	Sunset	Computerized	4	180	720
2	ELS - Math	Sunset	Computerized	4	180	720
2	ELS - Science	Sunset	Computerized	4	180	720
2	ELS - Social Studies	Sunset	Computerized	4	180	720
3	MKAS[2]	State	Computerized	1	50	50
3	STAR Reading	State	Computerized	3	20	60
3	PARCC ELA	State	Paper	2 (counting PBA)	75; 75; 60; 75	285
3	PARCC Math	State	Paper	2 (counting PBA)	75; 75; 75; 75	300
3	STAR Reading	Overton	Computerized	6	20	120
3	STAR Reading	Hillside	Computerized	3	20	60
3	STAR Math	Overton	Computerized	9	20	180
3	STAR Math	Hillside	Computerized	6	20	120
3	STAR Math	Mannequin	Computerized	3	20	60
3	STAR Math	Sunset	Computerized	3	20	60
3	i-Ready ELA	Overton	Computerized	3	60	180
3	i-Ready Math	Overton	Computerized	3	60	180
3	SRI	Hillside	Computerized	6	30	180
3	ELA Common Assessment	Mannequin	Paper	2	40-75	80–150
3	Math Common Assessment	Mannequin	Paper	2	40-75	80–150
3	ELS - ELA	Sunset	Computerized	4	180	720
3	ELS - Math	Sunset	Computerized	4	180	720
3	ELS - Science	Sunset	Computerized	4	180	720
3	ELS - Social Studies	Sunset	Computerized	4	180	720
4	PARCC ELA	State	Paper	2 (counting PBA)	75; 90; 60; 75	300
4	PARCC Math	State	Paper	2 (counting PBA)	80; 70; 75; 75	300
4	STAR Reading	Overton	Computerized	9	20	180
4	STAR Reading	Sunset	Computerized	3	20	60
4	STAR Math	Overton	Computerized	9	20	180

MASTER TABLE OF ASSESSMENTS, CONTINUED

Grade Level	Assessment	District or State Mandated?	Paper/Computerized	Test Frequency	Time for 1 Test (minutes)	Time for All Administrations
4	STAR Math	Sunset	Computerized	3	20	60
4	i-Ready ELA	Overton	Computerized	3	60	180
4	i-Ready ELA	Mannequin	Computerized	3	60	180
4	i-Ready Math	Overton	Computerized	3	60	180
4	i-Ready Math	Mannequin	Computerized	3	60	180
4	SRI	Hillside	Computerized	6	30	180
4	NWEA MAP ELA	Hillside	Computerized	2	60	120
4	NWEA MAP Math	Hillside	Computerized	2	60	120
4	ELA Common Assessment	Mannequin	Paper	2	40-75	80–150
4	Math Common Assessment	Mannequin	Paper	2	40-75	80–150
4	ELS - ELA	Sunset	Computerized	4	180	720
4	ELS - Math	Sunset	Computerized	4	180	720
4	ELS - Science	Sunset	Computerized	4	180	720
4	ELS - Social Studies	Sunset	Computerized	4	180	720
5	PARCC ELA	State	Paper	2 (counting PBA)	75; 90; 60; 75	300
5	PARCC Math	State	Paper	2 (counting PBA)	80; 70; 75; 75	300
5	Mississippi Science Test, 2nd Ed. (MST)	State	Computerized	1	180	180
5	STAR Reading	Overton	Computerized	9	20	180
5	STAR Reading	Sunset	Computerized	3	20	60
5	STAR Math	Overton	Computerized	9	20	180
5	STAR Math	Sunset	Computerized	3	20	60
5	i-Ready ELA	Overton	Computerized	3	60	180
5	i-Ready ELA	Mannequin	Computerized	3	60	180
5	i-Ready Math	Overton	Computerized	3	60	180
5	i-Ready Math	Mannequin	Computerized	3	60	180
5	Case21 Science	Overton	Computerized	2	90 - 150	180-300
5	NWEA MAP ELA	Hillside	Computerized	3	60	180
5	NWEA MAP Math	Hillside	Computerized	3	60	180
5	ELA Common Assessment	Mannequin	Paper	2	40-75	80–150
5	Math Common Assessment	Mannequin	Paper	2	40-75	80–150
5	ELS - ELA	Sunset	Computerized	4	180	720
5	ELS - Math	Sunset	Computerized	4	180	720

MASTER TABLE OF ASSESSMENTS, CONTINUED

APPENDIX C

Grade Level	Assessment	District or State Mandated?	Paper/Computerized	Test Frequency	Time for 1 Test (minutes)	Time for All Administrations
5	ELS - Science	Sunset	Computerized	4	180	720
5	ELS - Social Studies	Sunset	Computerized	4	180	720
6	PARCC ELA	State	Paper	2 (counting PBA)	75; 90; 60; 60; 60	345
6	PARCC Math	State	Paper	2 (counting PBA)	80; 70; 80; 75	305
6	STAR Reading	Overton	Computerized	9	20	180
6	STAR Reading	Sunset	Computerized	3	20	60
6	STAR Math	Overton	Computerized	9	20	180
6	i-Ready ELA	Overton	Computerized	3	60	180
6	i-Ready ELA	Mannequin	Computerized	3	60	180
6	i-Ready Math	Overton	Computerized	3	60	180
6	i-Ready Math	Mannequin	Computerized	3	60	180
6	NWEA MAP ELA	Hillside	Computerized	3	60	180
6	NWEA MAP Math	Hillside	Computerized	3	60	180
6	Orleans-Hanna	Hillside	Paper	1	60	60
6	ELA Common Assessment	Mannequin	Paper	2	40-75	80-150
6	Math Common Assessment	Mannequin	Paper	2	40-75	80-150
6	ELS - ELA	Sunset	Paper	4	180	720
6	ELS - Math	Sunset	Paper	4	180	720
6	ELS - Science	Sunset	Paper	4	180	720
6	ELS - Social Studies	Sunset	Paper	4	180	720
7	PARCC ELA	State	Paper	2 (counting PBA)	75; 90; 60; 60; 60	345
7	PARCC Math	State	Paper	2 (counting PBA)	80; 70; 80; 75	305
7	i-Ready ELA	Overton	Computerized	3	60	180
7	i-Ready ELA	Mannequin	Computerized	3	60	180
7	i-Ready Math	Overton	Computerized	3	60	180
7	i-Ready Math	Mannequin	Computerized	3	60	180
7	STAR Reading	Overton	Computerized	9	20	180
7	STAR Reading	Sunset	Computerized	3	20	60
7	STAR Math	Overton	Computerized	9	20	180
7	NWEA MAP ELA	Hillside	Computerized	3	60	180
7	NWEA MAP Math	Hillside	Computerized	3	60	180
7	ELA Common Assessment	Mannequin	Paper	2	40-75	80-150
7	Math Common Assessment	Mannequin	Paper	2	40-75	80-150

MASTER TABLE OF ASSESSMENTS, CONTINUED

Grade Level	Assessment	District or State Mandated?	Paper/Computerized	Test Frequency	Time for 1 Test (minutes)	Time for All Administrations
7	ELS - ELA	Sunset	Paper	4	180	720
7	ELS - Math	Sunset	Paper	4	180	720
7	ELS - Science	Sunset	Paper	4	180	720
7	ELS - Social Studies	Sunset	Paper	4	180	720
8	PARCC ELA	State	Paper	2 (counting PBA)	75; 90; 60; 60; 60	345
8	PARCC Math	State	Paper	2 (counting PBA)	80; 70; 80; 75	305
8	Mississippi Science Test, 2nd Ed. (MST)	State	Computerized	1	180	180
8	i-Ready ELA	Overton	Computerized	3	60	180
8	i-Ready ELA	Mannequin	Computerized	3	60	180
8	i-Ready Math	Overton	Computerized	3	60	180
8	i-Ready Math	Mannequin	Computerized	3	60	180
8	STAR Reading	Overton	Computerized	9	20	180
8	STAR Reading	Sunset	Computerized	3	20	60
8	STAR Math	Overton	Computerized	9	20	180
8	Case21 Science	Overton	Computerized	2	90 - 150	180-300
8	NWEA MAP ELA	Hillside	Computerized	3	60	180
8	NWEA MAP Math	Hillside	Computerized	3	60	180
8	Curriculum ELA	Hillside	Paper	1	70; 70	140
8	Curriculum Math	Hillside	Paper	1	60; 60	120
8	Curriculum Science	Hillside	Paper	1	45; 45	90
8	ELA Common Assessment	Mannequin	Paper	2	40-75	80-150
8	Math Common Assessment	Mannequin	Paper	2	40-75	80-150
8	ELS - ELA	Sunset	Paper	4	180	720
8	ELS - Math	Sunset	Paper	4	180	720
8	ELS - Science	Sunset	Paper	4	180	720
8	ELS - Social Studies	Sunset	Paper	4	180	720
9	PARCC Algebra I	State	Computerized	2 (counting PBA)	90; 75; 90; 75	320
9	SATP2 Biology I	State	Computerized	1	120-240	120-240
9	Curriculum ELA I	Hillside	Paper	1	120	120
10	Curriculum Math II	Hillside	Paper	1	90; 150	240
9	Curriculum Biology	Hillside	Paper	1	75; 45	120
9	Curriculum World History	Hillside	Paper	1	120; 90	210
9	English I Common Assessment	Mannequin	Paper	1	50-75	50-75

MASTER TABLE OF ASSESSMENTS, CONTINUED

Grade Level	Assessment	District or State Mandated?	Paper/Computerized	Test Frequency	Time for 1 Test (minutes)	Time for All Administrations
9	Algebra I Common Assessment	Mannequin	Paper	2	50-75	100-150
9	STAR Reading	Sunset	Computerized	3	20	60
9	ELS - ELA	Sunset	Paper	4	180	720
9	ELS - Math	Sunset	Paper	4	180	720
9	ELS - Science	Sunset	Paper	4	180	720
9	ELS - Social Studies	Sunset	Paper	4	180	720
10	PARCC English II	State	Computerized	2 (counting PBA)	75; 90; 60; 60; 60	345
10	Curriculum ELA II	Hillside	Paper	1	90; 45	135
10	Curriculum Chemistry	Hillside	Paper	1	75; 45	120
10	English II Common Assessment	Mannequin	Paper	1	50-75	50-75
10	STAR Reading	Sunset	Computerized	3	20	60
10	ELS - ELA	Sunset	Paper	4	180	720
10	ELS - Math	Sunset	Paper	4	180	720
10	ELS - Science	Sunset	Paper	4	180	720
10	ELS - Social Studies	Sunset	Paper	4	180	720
11	SATP2 US History	State	Computerized	1	120-240	120-240
11	ACT	State	Paper	1	175	175
10	Curriculum American History	Hillside	Paper	1	120; 120	240
11	STAR Reading	Sunset	Computerized	3	20	60
11	ELS - ELA	Sunset	Paper	4	180	720
11	ELS - Math	Sunset	Paper	4	180	720
11	ELS - Science	Sunset	Paper	4	180	720
11	ELS - Social Studies	Sunset	Paper	4	180	720
12	STAR Reading	Sunset	Computerized	3	20	60
12	ELS - ELA	Sunset	Paper	4	180	720
12	ELS - Math	Sunset	Paper	4	180	720
12	ELS - Science	Sunset	Paper	4	180	720
12	ELS - Social Studies	Sunset	Paper	4	180	720

APPENDIX D: STANDARDIZED TESTS IN _____ DISTRICT

[The district should complete one page (front and back if necessary) for each grade level.]

Grade Level	Description of State Tests Required this Year	Length of One Administration (Minutes)	Frequency	Month(s) Administered
[GRADE LEVEL] **[Specify the group of students who take the test]**	• What is the name of the test? (Include all names of the test.) • What does the test measure? • How are the results used?	How long is the test? Specify each part.	How frequently is the test administered?	When will students take the test?
3rd Grade (All Students) **(Special exemptions apply for students with special needs.)**	**The MKAS^2 test** measures reading proficiency in the fall and spring of kindergarten and third grade. If third graders do not receive a passing score on the MAAP ELA exam, then schools can use a passing score on the MKAS^2 as an alternative. Students who do not pass the MAAP ELA exam will receive targeted reading interventions and be given two more chances to pass MKAS^2.	50	1	August, May
3rd Grade (All Students) **(Accommodations and modifications are provided according to the student's IEP.)**	**The STAR Reading test** measures students' reading ability by identifying their skill level and growth. The STAR Reading test helps teachers identify students who need reading interventions and what types. The results are used to ensure that students are making appropriate progress to pass the MAAP ELA exam. The state requires the test to be given at least three times.	~20	3	September, December, May
3rd Grade (All Students) **(Accommodations and modifications are provided according to the student's IEP.)**	**The MAAP Math Exam** measures student progress in math according to the Mississippi College and Career Readiness Standards. The Mississippi College and Career Readiness Standards describe what all students should know and be able to do at the end of each grade level. The results are used to calculate our accountability rating and inform parents of their children's end-of-year progress.	127 (operational test); 70 (performance task) 197 TOTAL	1	May
3rd Grade (All Students) **(Accommodations and modifications are provided according to the student's IEP.)**	**The MAAP ELA Exam** measures student progress in English and Language Arts according to the Mississippi College and Career Readiness Standards. The Mississippi College and Career Readiness Standards describe what all students should know and be able to do at the end of each grade level. The results are used to calculate our accountability rating and inform parents of their children's end-of-year progress. In third grade, the ELA exam is also the primary test used for the Third Grade Reading Gate. Every third grader in the state is required to receive a passing score on the test, which is a score of passing, proficient, or advanced, or a passing score on the MKAS^2 in order to advance to the next grade.	113 (operational test); 150 (performance task) 263 TOTAL	1	May

Grade Level	Description of District Tests Required this Year	Length of One Administration (Minutes)	Frequency	Month(s) Administered
[GRADE LEVEL] **[Specify the group of students who take the test]**	• What is the name of the test? (Include all names of the test.) • What does the test measure? • How are the results used?	How long is the test? Specify each part.	How frequently is the test administered?	When will students take the test?
3rd Grade (All Students) **(Accommodations and modifications are provided according to the student's IEP.)**	**The STAR Reading test** measures students' reading ability by identifying their skill level and growth. The STAR Reading test helps teachers identify if students need reading intervention, and it tells teachers what types of interventions students need. The results are used to ensure that students are making appropriate progress to pass the MAAP ELA test. The district requires the test to be given an additional three times, for a total of six times including the three state administrations.	~20	3	October, February, March
3rd Grade (All Students) **(Accommodations and modifications are provided according to the student's IEP.)**	**The STAR Math test** measures students' math performance. The STAR Math test helps teachers identify what math concepts students understand and which ones they are struggling with. The results are used to help teachers plan targeted interventions for students.	~20	6	September, October, December, February, March, May

Learn More

The executive summary and full report can also be found online at **mississippifirst.org**. Mississippi First is a nonpartisan, nonprofit that champions transformative policy solutions ensuring educational excellence for every Mississippi child.

MISSISSIPPI FIRST