

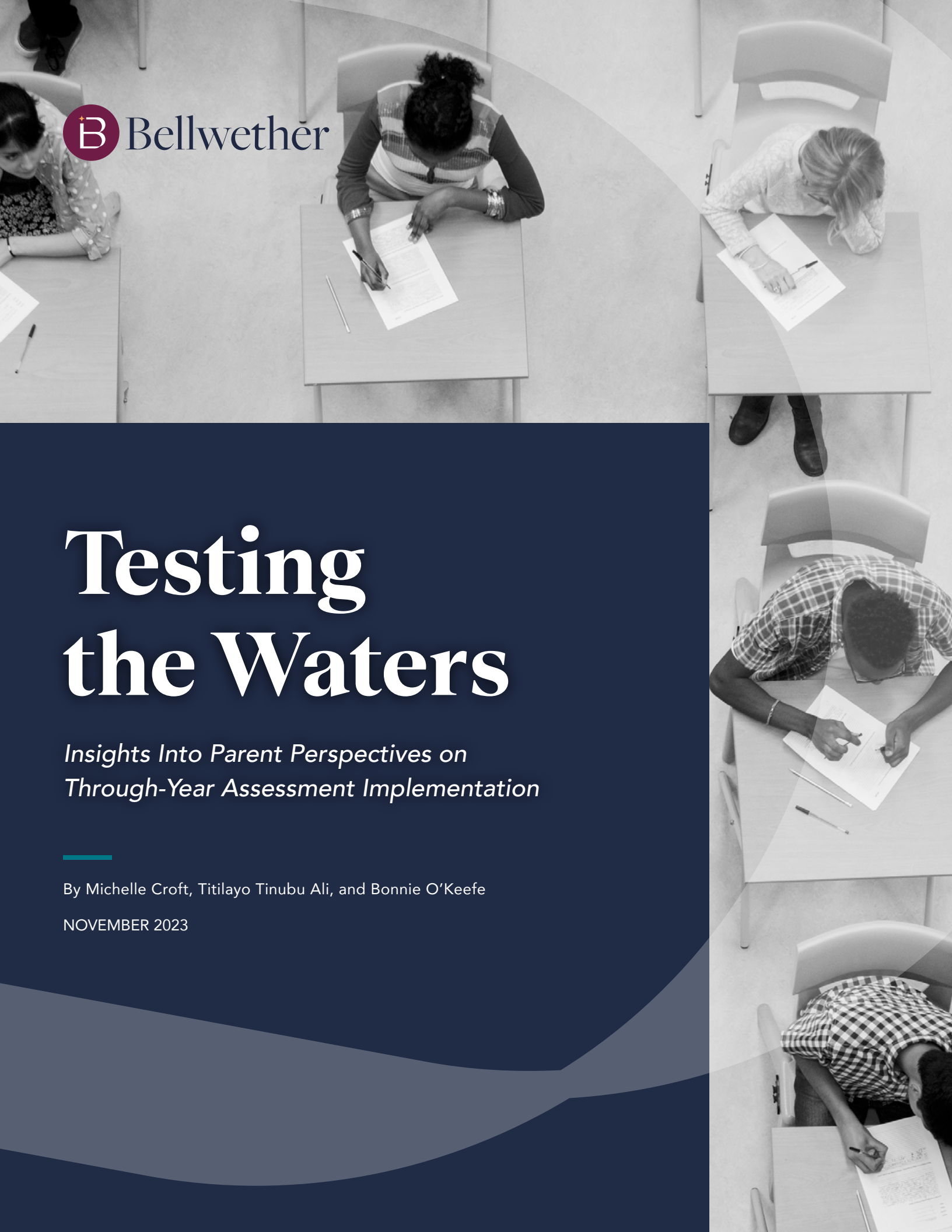


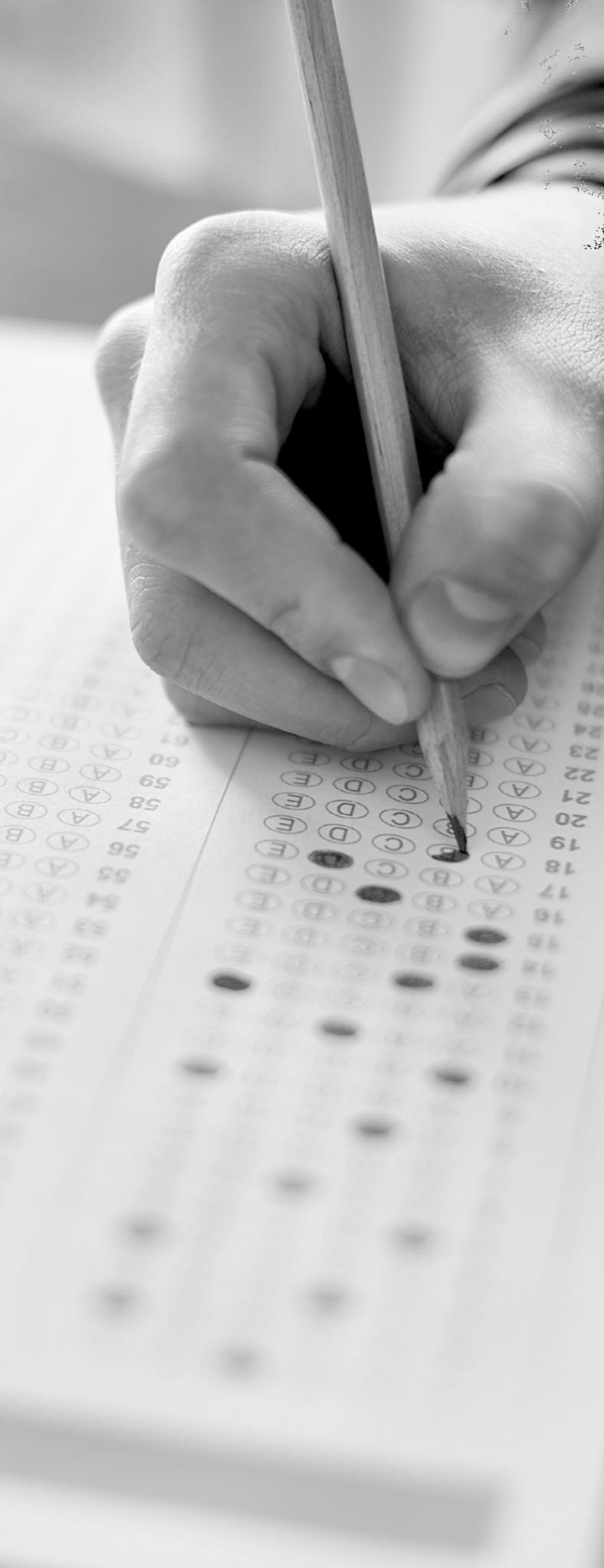
Testing the Waters

*Insights Into Parent Perspectives on
Through-Year Assessment Implementation*

By Michelle Croft, Titilayo Tinubu Ali, and Bonnie O'Keefe

NOVEMBER 2023





| CONTENTS

- 3 EXECUTIVE SUMMARY
- 5 INTRODUCTION
- 6 WHAT ARE STATES REQUIRED TO TEST?
- 8 HOW STATES ARE DESIGNING THROUGH-YEAR ASSESSMENTS
- 10 METHODOLOGY
- 12 STATE PROFILES
 - Delaware
 - Florida
 - Nebraska
 - Texas
- 17 AREAS OF OPPORTUNITY
- 26 WHY NOT LEAVE INTERIM ASSESSMENTS TO SCHOOL DISTRICTS?
- 27 RECOMMENDATIONS
- 30 CONCLUSION
- 31 APPENDIX
- 32 ENDNOTES
- 35 ACKNOWLEDGMENTS
 - ABOUT THE AUTHORS
 - ABOUT BELLWETHER

Executive Summary

Statewide assessments are an important tool to monitor K-12 students' learning at the end of each school year and to share that information with educators, families, and communities.

But what if instead of a single, longer test at the end of the year, with scores trickling in during the summer or fall, state tests were shorter and embedded throughout the school year? Several states are implementing or exploring this idea of “through-year” assessments, where multiple tests are administered over a school year.¹

Many technical, test design, and implementation questions are involved in designing new statewide assessments. These are all critical, shaping often underexamined aspects of testing: student test experience and families' perceptions of testing in this new structure. Although the state outreach to students and families is needed generally, it is especially vital to elevate the experiences and perspectives of historically marginalized students and families, such as students of color, students with disabilities, and English learners (ELs), to understand how new test designs can best meet the needs of all students and help close inequitable achievement gaps.

This report identifies four states that have taken different approaches in designing new through-year assessments: Delaware, Florida, Nebraska, and Texas. In each state, we spoke with state assessment directors and district leaders to learn more about how the tests worked for students, families, and educators. In all states except Texas, which is in the early stages of its pilot, we also spoke with Black and Latino parents and parents of students with disabilities to hear about their experiences with the new assessments.

These conversations surface five opportunities for improving through-year assessment implementation:

- Incorporating parent voices in assessment design and implementation decisions.
- Improving information flow between the state and parents, via districts and schools.

- Monitoring of student test experience and the effects of new test administrations on students' school day and year.
- Reporting of scores with more explanatory and framing information on what new scores signify.
- Clarifying the use cases of different test designs for accountability versus instructional decisions.

Five policy recommendations emerge for states interested in improving implementation or other kinds of assessment innovation:

- 1. Design new assessment systems to solve an agreed-upon problem identified by districts and their communities.** Through-year assessments are more effective when districts and communities share early buy-in on the model or when it is a district- or community-led initiative to create the new assessment model.
- 2. Increase family and community engagement at multiple steps in the assessment design process.** Most parents we spoke with felt like they did not have a voice regarding assessments, and they would like to have one. Developing stronger communication channels with parents can help states better tailor the assessment to the needs of the community and avoid misunderstandings or distrust of the assessment and the scores.

3. Provide districts with timely, robust resources to implement and disseminate information about new tests. Districts largely rely on state resources for communication. Districts need timely, robust resources to assure correct implementation of the testing program and communicate changes in test administration, uses, and scoring with their teachers, families, and communities.

4. Implement formal continuous improvement processes based on monitoring of student experiences. Given that through-year assessments are new, states and districts should make a concerted effort to identify unintended consequences related to the tests by more thoroughly monitoring and evaluating testing.

5. Explore making changes to federal law to achieve the stated goals of through-year assessments.

The goals of through-year assessments are to reduce testing time; provide more timely, instructionally relevant results; and use the results for accountability. These goals magnify the tension between the accountability and instructional uses of the test scores. Some of the tension is codified in the Every Student Succeeds Act (ESSA), the federal law requiring state assessments, and may require changes to the law to achieve the goals.²

It is especially vital to elevate the experiences and perspectives of historically marginalized students and families, such as students of color, students with disabilities, and English learners.



Introduction

Each spring, students in grades 3-8 and high school take their state’s end-of-year test. These tests, required by ESSA, are used for a variety of purposes, such as directing resources to struggling schools and informing parents and the general public about how well a school is educating all students.³ Because of their end-of-year timing, the tests cannot help educators and families know if a student is off-track until it is too late to take action.

Some states are rethinking their approach to these state assessments to better measure student progress within the school year, identify students early in the year who need additional support, and shorten testing times. New through-year assessments are multiple test events administered over the course of a school year “as part of an assessment system designed to produce a single summative score meeting federal and state accountability requirements.”⁴

As these assessments are new, much is unknown about how they impact students — either positively or negatively. There is also quite a bit of variation in how states design the assessments, so one state’s through-year assessment may impact students very differently than that of another state’s through-year assessment. We use examples from four states — **Delaware, Florida, Nebraska, and Texas** — to discuss how the design features in all stages of testing, from test development to test day to receiving and using the test results, impact students’ testing experience. In particular, we focus on the experiences of parents and students from historically marginalized backgrounds, specifically Black and Latino students and students with disabilities.⁵

Parents are important stakeholders when it comes to student assessment, as state assessments are one of the only mechanisms that allow parents to see how their child is doing academically compared to their

peers within the state.⁶ Testing and score reports have been criticized historically for failing to communicate information to parents⁷ and, given the nuances and complexities of understanding the scores from through-year assessments, this problem may be exacerbated.

Finally, in any new education policy endeavor, it is critical to consider its effects on students and families furthest from educational opportunity and design solutions through a focus on the needs, contexts, behaviors, and emotions of the people that the solutions will serve.⁸ Students and families have lived experiences of the problems that through-year assessments aim to solve, and they may bring new perspectives, challenge existing approaches, and offer creative solutions. For families of color and parents of students with disabilities, the ability to contribute to implementation decision-making and better understand through-year assessments and scores is particularly important. These student groups are traditionally underserved in schools;⁹ through-year assessments provide an opportunity to identify students in need of additional supports earlier in the school year and tighten communication among families, students, and teachers. If the needs of these families and students are not explicitly prioritized in policy design and test implementation, there is a risk that solutions may not match the need and may have unintended consequences that disproportionately impact them.

What Are States Required to Test?

Under ESSA, states are required to test students annually in grades 3-8 and high school for reading and math as well as science for certain grade spans.¹⁰ States may administer additional tests in other content areas (e.g., social studies) or grades (e.g., grade 2). The tests are designed to measure the information students were supposed to have been taught during the school year, based on the state's content standards for each grade.¹¹

States must publicly report scores overall and for certain subgroups of students, such as by race and disability status.¹² The scores help inform parents and the broader public about how well schools are serving students and are used by states to direct resources or supports to schools.

States, districts, and schools also use scores in other ways outside of federal accountability. For example, some states use the scores as a component of teacher evaluation¹³ and school ratings outside of ESSA.¹⁴ Similarly, for charter schools, the scores can inform decisions about school closures.¹⁵

In addition to accountability, assessment scores are used for instructional purposes, some of which were intended by ESSA. For example, teachers use the scores to better understand student performance, and administrators use them to track progress and understand whether interventions (e.g., implementation of a reading curriculum) are working.¹⁶ Schools may also use the scores, in conjunction with other information about students, to place them in certain programs, either for advancement or remediation.¹⁷

Why Are States Considering Through-Year Assessments?

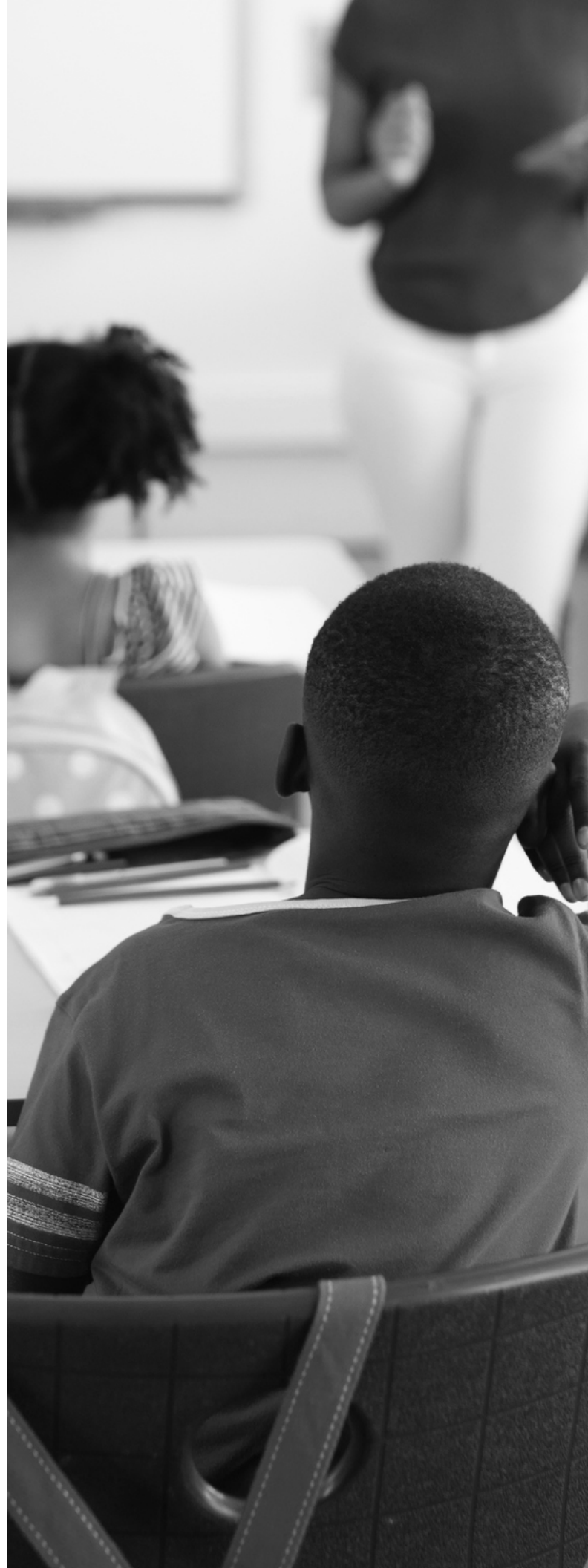
Historically, most states have opted to use one end-of-year assessment (hence, "summative testing") to measure how much a student has learned within the year. However, there are limitations to having one end-of-year test, causing states to rethink that strategy. Four of the most frequently cited limitations of one end-of-year assessment include:

- 1. Late results.** Because many states include writing or open-ended items that must be scored, the scores are often not available until the summer or the start of the next school year. Some refer to these tests as a "post-mortem" because students are in the next grade by the time results are available and it is too late to adapt instruction.
- 2. Lack of actionable information for educators and families.** Typical end-of-year assessments are good for providing higher-level information about how schools, districts, and other systems are educating students, but they cannot provide the level of detailed information needed to directly aid day-to-day instruction or provide parents with timely information on their individual child's learning and growth.
- 3. Long testing times.** Educators, families, and school administrators have long complained about the length of end-of-year state tests.¹⁸ In recent years, states have reduced the testing time, but they continue to disrupt whole days or weeks of school.¹⁹ This is often caused by the logistical necessities of administering multiple, secure tests on computers to most or all students in a school. Additionally, districts often supplement state tests with their own assessments to inform instruction during the year, multiplying the total time spent on testing for students and teachers.

4. Constraints on growth measures. Growth measures (i.e., how much progress a student makes over a period of time) are critical to measuring and incentivizing student progress and are a central part of many states' accountability systems. But by only measuring growth from one school year to the next, states must grapple with missing data for mobile students or unusual course progressions, summer learning loss, and mismatches in standards between grade levels.²⁰

These goals of faster reporting, less administrative time, more instructional relevance within the school year, and newer growth measures have led nearly a dozen states to adopt — or consider adopting — through-year assessments.²¹

Historically, most states have opted to use one end-of-year assessment (hence, “summative testing”) to measure how much a student has learned within the year. However, there are limitations to having one end-of-year test, causing states to rethink that strategy.



How States Are Designing Through-Year Assessments

For states investigating through-year assessments,²² ESSA allows them to use multiple tests administered throughout the year as long as the assessment system meets certain requirements, such as adequately measuring the state's content standards at a particular grade level.²³ Beyond the ESSA requirements, there are multiple design decisions ranging from when the assessments should be administered to what will be tested and whether to aggregate the scores.²⁴

PRODUCING A SUMMATIVE SCORE

Although states typically administer two to three tests throughout the year, many through-year assessments still use only the last test (sometimes called a "test event") to produce an end-of-year score for accountability, instead of combining performance information from multiple points throughout the year. For example, states are only using the students' spring test score, instead of using a combination of the students' fall, winter, and spring scores. For purposes of this report, we are focusing on state tests that either are or could produce a single summative score, because through-year assessments are still "a fundamental shift" in conceptualizing a statewide assessment system.²⁵

From our conversations with officials in Delaware, Florida, Nebraska, and Texas, the most significant design decision is how the assessment aligns with standards and curriculum.

Under ESSA, all states must design a test that measures what students were supposed to be taught that year, which is called aligning the test to the standards.²⁶ If a state adopts a through-year model, it begins to have flexibility in *when* those standards are assessed.

In our conversations with states, this flexibility in when to assess particular standards through the connection to the curriculum was named the most significant design feature.²⁷ Historically, most states have been averse to implicitly or explicitly directing scope, sequence, or curriculum via testing, and they leave it to districts to interpret standards and choose or create a curriculum.²⁸

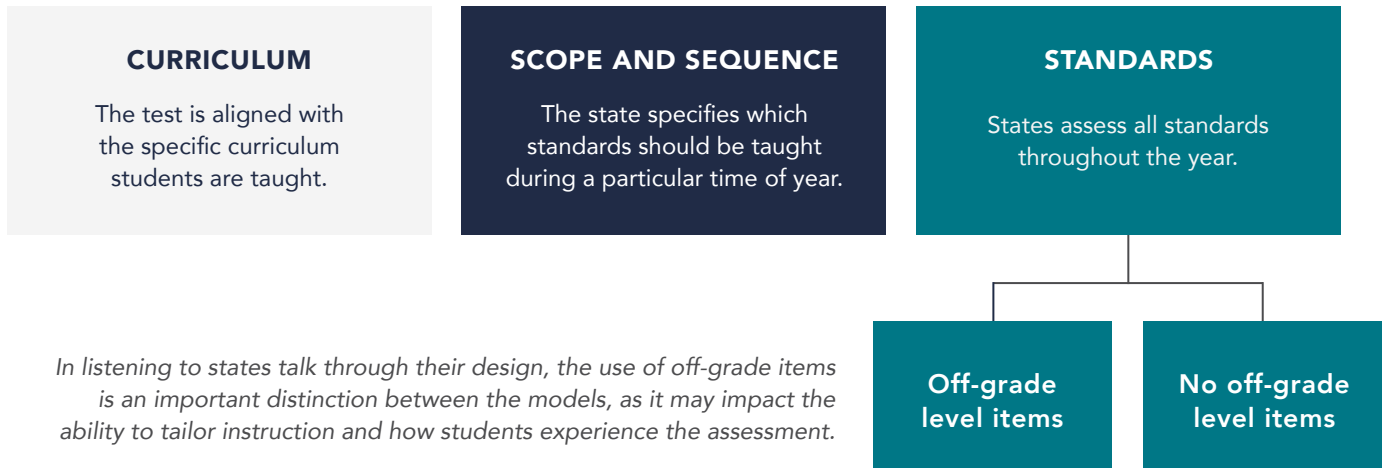
But more states are getting active in curriculum quality and adoption in the following ways:²⁹

- **Curriculum-aligned:** The test is aligned with the specific curriculum students are taught.
- **Scope- and sequence-aligned:** The state specifies which standards should be taught and tested during a particular time of year.
- **Standards-aligned:** States assess all standards throughout the year.

Within standards-aligned adoption, the decision of whether or not to include off-grade items is another significant design element impacting the way test scores can be used and understood (Figure).

Because these assessments are still in their early stages, there may be other types of design considerations, such as the ability to roll up scores from earlier tests into a composite score, that will become increasingly important or introduce new variations among state through-year assessments.

FIGURE: RELATIONSHIP AMONG CURRICULUM, SCOPE AND SEQUENCE, AND STANDARDS



If a state adopts a through-year model, it begins to have flexibility in when those standards are assessed. In our conversations with states, this flexibility in *when* to assess particular standards through the connection to the curriculum was named the most significant design feature.

Methodology

To better understand how the different assessment models impact historically marginalized students' testing experience, we used data from Education First's report "Through-Year Curriculum-Connected Assessment," which details the through-year assessment landscape for the 2022-23 school year. We used the data to identify states representing the different assessment models, prioritizing those that were closer to statewide operation.³⁰ We also tried to balance the selection of states with sufficient diversity (geography and student populations).

Based on the landscapes, we selected four states — Delaware, Florida, Nebraska, and Texas — representing three of the four models (Figure).

- **Standards-aligned, on-grade-level items:** The Florida Assessment of Student Thinking (FAST) and Texas' Through-Year Assessment Pilot. Texas was selected because it has the potential to impact large numbers of students. Since the state is in the early stages of its pilot, parents are less aware of the assessment, so we do not focus as heavily on Texas as on the other states.
- **Scope- and sequence-aligned assessment:** Delaware's Social Studies Through-Year Assessment.
- **Standards-aligned, includes off-grade-level items:** The Nebraska Student-Centered Assessment System (NSCAS) Growth assessment.

Curriculum-aligned assessments are not examined in this report because this model is currently in review under Louisiana's participation in the Innovative Assessment Demonstration Authority, and other states are less likely to adopt this model of assessment.³¹ Given the history of local control in education, many states do not have the capability (either legally or through political capital) to define a scope and sequence or require a particular curriculum.

To better understand the state assessments and how they impact students, particularly Black and Latino students and students with disabilities, we:

- **Interviewed state assessment directors and 14 school or district leaders in each state (Delaware, Florida, Nebraska, and Texas).** Of district leaders, we primarily spoke with assessment directors, as well as instructional leaders in Nebraska and teacher leaders of the Social Studies Coalition of Delaware.
- **Consulted with schools and community groups who serve families within each state.** Many of the community groups did not work directly with assessments but were willing to share information about the project with the families they served.
- **Spoke with 20 parents in Delaware, Florida, and Nebraska, through either interviews or focus groups.** The parents we spoke with either identified as a person of color and/or they had a child with a disability receiving accommodations on tests.³² We particularly wanted to hear from these communities since they often do not have input into assessment decisions. We spoke with the greatest number of parents in Florida, which was the furthest along in implementation, but we also spoke with parents in Delaware and Nebraska (Appendix). Because Texas is in the early stages of its pilot and districts reported not sharing testing information with parents, we did not speak with parents in that state. We also collected survey data from nearly 400 Florida parents, which we predominantly used to identify interviewees and gather high-level information on parents' perceptions of their child's test experience.



There are some limitations to our report. Although we did receive diverse perspectives through the interviews, the parent perspectives should be interpreted as illustrative and not representative. We intentionally chose a purposive sample of parents at this stage in the development and implementation of through-year state assessments to focus most on certain subgroups of students (e.g., Black and Latino students and students with disabilities) and go deeper in interviews rather than seeking a broad polling-style sample. Also, given the ages of the students primarily taking these tests (grades 3-8), we initially explored the feasibility of connecting directly with students but ultimately chose to rely on information from parents and district officials instead.

Another limitation is our ability to recruit monolingual, Spanish-speaking parents. Although we made concerted attempts to recruit such parents by translating all recruitment materials into Spanish and reaching out to organizations serving Spanish-speaking parents, our efforts were not successful within our research timeline. We did speak with at least one parent who was Spanish-speaking, bilingual who preferred to participate in the interview using English. This gap in parent perspectives is an area where future research is needed.³³

Given the history of local control in education, many states do not have the capability (either legally or through political capital) to define a scope and sequence or require a particular curriculum.

State Profiles

Through-year assessments vary from state to state. For the four states analyzed in this report — **Delaware, Florida, Nebraska, and Texas** — profiles below provide insight into the through-year assessment landscape in each region (Table).

TABLE: THROUGH-YEAR ASSESSMENT STATE PROFILES (2022-23 SCHOOL YEAR)

State	Grade Levels and Subject Areas	School Year Started	Implementation Status	Type of Through-Year Assessment
Delaware	<ul style="list-style-type: none"> Grade 4 civics and history Grade 5 economics and geography Grade 6 geography Grade 7 civics and economics Grade 8 history 	Varies by subject; grade 8 history was the first to pilot in 2021-22	<ul style="list-style-type: none"> Grade 8 history is operational Grade 4 and grade 6 were piloted during the 2022-23 school year and are operational in the 2023-24 school year Grade 5 and grade 7 will be piloted during the 2023-24 school year 	Scope- and sequence-aligned
Florida	<ul style="list-style-type: none"> Pre-K through grade 10 English language arts (ELA) Pre-K through grade 8 math 	2022-23	Operational statewide	Standards-aligned, only on-grade-level items
Nebraska	<ul style="list-style-type: none"> Grades 3-8 math and ELA 	2020-21	Operational statewide, with districts selecting whether or not to participate in the fall and winter test events	Standards-aligned, includes off-grade-level items that do not contribute to a student's proficiency score
Texas	<ul style="list-style-type: none"> Grade 5 science Grade 6-7 math Grade 8 social studies 	2022-23	Voluntary pilot, approximately 8-10% of school districts and charters opted to participate	Standards-aligned, only on-grade-level items

State Profile: Delaware

Background

The Delaware Social Studies Standards are designed for grade bands instead of grade levels. Schools are required to teach four content areas — Civics, Economics, Geography, and History — over the course of two years for students while they are in grades 4-5 and within three years for students in grades 6-8.

Districts and schools have traditionally struggled to map the four social studies content areas within the two- to three-year grade clusters. This has resulted in a wide array of approaches to content and course design.

The Social Studies Coalition, a network of the majority of Delaware local education agencies, partnered with the Delaware Department of Education to articulate the standards into a new course design with the development of recommended curricular materials so that it was clearer *what* content should be taught *when*. The change helped to add specificity regarding consistent expectations and experiences across the state. The partnership also contributed to the design of the through-year social studies assessment system.

Grade Levels and Subject Areas

Delaware's phasing in of through-year assessments is piloted in the initial year and operational in the second year of implementation.³⁴ Starting in the 2021-22 school year, the state implemented new recommended curriculum in eighth-grade history and piloted the through-year assessment. The eighth-grade history assessment was operational for the 2022-23 school year. Also in the 2022-23 school year, the state piloted fourth-grade civics and history and sixth-grade geography assessments. Delaware is piloting fifth grade and seventh grade in the current 2023-24 school year.

Test Events

The state administers three test events.³⁵ They include fall (November to early December), winter (mid-February to early March), and spring (May). Each testing window is approximately three weeks.³⁶

Test Length

The test is untimed. Students may take it over two days if necessary; however, it is designed to be administered within one class period.³⁷

Score Reporting

Students receive their raw score (i.e., the number answered correctly) immediately after submitting their test.

Although eighth-grade scores are included in the test portal, the state did not provide individual score reports during the 2022-23 school year because it was waiting to share score reports with parents after the summer standard-setting process.

How Each of the Individual Tests May Be Used for Accountability Purposes

The scores are used as part of Delaware's accountability system. To calculate the score, the state uses a gradebook method in which each test is rolled up into a total score.

State Profile: Florida

Background

Gov. Ron DeSantis championed the state's FAST assessment to:

- Shift from the end-of-year assessments to “progress monitoring for accountability.”³⁸
- Minimize the stress of testing by having three shorter tests instead of one longer end-of-year assessment.
- Reduce testing time.

The Florida Legislature codified the proposal in Senate Bill 1048 (2022).³⁹

The assessments are adaptive in that a student's performance on the first progress-monitoring tool helps inform the starting point on the subsequent assessment.

Grade Levels and Subject Areas

Pre-K through grade 10 for ELA and pre-K through grade 8 for math.⁴⁰

Test Events

FAST is administered three times a year, and each administration is called a “PM” (short for progress monitoring).⁴¹

PM1 occurs at the beginning of the year from mid-August through the end of September. PM2 is the midyear test event, taking place in early December through the end of January. PM3 is the end-of-year test, which is administered in May through early June.⁴²

Test Length

The recommended testing time for PM1 and PM2 varies based on grade and content area and ranges from 80 to 100 minutes. PM3 ranges from 100 to 120 minutes.⁴³

Score Reporting

Parents can see their child's score in the test portal 24 hours after the test is completed.

How Each of the Individual Tests May Be Used for Accountability Purposes

Currently, only the score for PM3 is used for accountability. As part of Senate Bill 1048, the state will be studying whether PM1 and PM2 can be used for accountability purposes.

State Profile: Nebraska

Background

Nebraska's NSCAS Growth assessment grew out of an interest from the state and some districts to reduce the amount of testing and provide faster results.

The state wanted the assessments to have instructionally useful information and created a test in two parts. The first part of the test is on-grade-level (i.e., questions designed to measure the state's standards for the student's current grade). The second part of the test adapts to provide off-grade-level items (i.e., questions designed to measure the state's standards for the grade above or below the student's current grade) so that educators may have a better sense of students' areas of strengths and growth. Information from the earlier test events also helps to identify where students start on later test events.⁴⁴

Districts are currently required to participate in the spring test event but may participate in fall and winter testing. The state had planned to require districts to implement the fall test in the 2022-23 school year⁴⁵ but delayed the change to require the fall test event.⁴⁶ In January 2023, the State Board of Education voted to postpone moving to a through-year model for the 2023-24 school year,⁴⁷ such that the fall and winter test events are optional in the 2023-24 school year.⁴⁸ The delays were due, in part, to technical concerns about the assessment, such as the amount of error associated with the test score and discrepancies between the NSCAS Growth scores and the district-administered interims, as well as making improvements to the assessment, such as improving score reporting.⁴⁹

In response to concerns from districts, the State Department of Education also made changes to the types of off-grade-level items a student may see on the test. Originally, students would only receive a grade level higher or lower than their enrolled grade. For the 2022-23 school year, the State Department of Education adjusted the constraints on the computer-adaptive test

engine, which selects the items a student will see based on their previous answers, so that students could receive up to two grade levels above and unlimited grade levels below their enrolled grade level.

Grade Levels and Subject Areas

Grades 3-8 math and ELA.

Test Events

The state offers three test events, although districts do not have to participate in both the fall and winter test event.⁵⁰ All must participate in the spring test event. The fall test is administered from mid-August until the end of September. The winter test may be administered from early December until the end of January. The spring test is administered in April to mid-May.

Test Length

The state recommends a 90-minute testing time for math and ELA, and there are approximately 45 questions.⁵¹

Score Reporting

Schools may download score reports to share with parents.

How Each of the Individual Tests May Be Used for Accountability Purposes

Originally, the state sought to bank performance on individual items, but it shifted to planning to bank performance on test events. For example, if a student met the proficiency standard during the fall test, the student could use that score for end-of-year proficiency.

In 2023, the state determined that only the end-of-year test event could be used. The decision was in part because the test design could not guarantee that a student would meet the on-grade-level blueprint for the fall and winter assessment and in part because the state interpreted ESSA as requiring the proficiency decision to be based on end-of-year knowledge and skills.

State Profile: Texas

Background

The Texas Through-Year Assessment Pilot is a legislatively required pilot program under House Bill 3906.⁵² The bill identified goals that the statewide assessments are as short as practicable and minimize the disruption to the educational program. The pilot explores whether a through-year assessment model, which provides students multiple opportunities throughout the school year to demonstrate their mastery of standards, is a possible replacement for the state's current end-of-year test.

House Bill 3906 is fairly vague and allowed the Texas Education Agency (TEA) flexibility in designing the pilot. The TEA conducted stakeholder focus groups to help inform the design of the assessment (Disclosure).

The test uses an adaptive design so that how the student performs on earlier test events helps identify where they will start on later test events.⁵³

Grade Levels and Subject Areas

For the 2022-23 school year, the pilot was available in grade 5 science, grades 6-7 math, and grade 8 social studies.⁵⁴

Test Events

The state administers three test events. They include Opportunity 1 (mid-November), Opportunity 2 (end of January/beginning of February), and Opportunity 3 (end of March/early April).⁵⁵ Each test window is approximately one week.

Test Length

For Opportunities 1 and 2, the tests are designed to be as short as possible, with 17-20 questions per test.

The state planned for Opportunity 3 to be longer than those in Opportunities 1 and 2 but still be shorter than the end-of-year State of Texas Assessments of Academic Readiness.⁵⁶

Score Reporting

Because the test is in the pilot stage, the scores are not included in the state's test portal. Teachers can access the scores and print out individual score reports for families.

How Each of the Individual Tests May Be Used for Accountability Purposes

The scores are not currently used for accountability. The TEA plans to use the pilot data collected over multiple years to determine if it is possible to roll up the individual scores into a summative score.⁵⁷

Areas of Opportunity

Through conversations with state, district, and community organization leaders and parents, five areas of opportunity emerged. The opportunities relate to communication and the need for additional resources but also to how the tests impact a student's educational experience. Some areas of opportunity are unique to through-year assessment models, and others are areas where state assessments more generally can improve when switching to a new assessment. Areas of opportunity include:

1. Incorporating parent voices in assessment design and implementation decisions.
2. Improving information flow between the state and parents, via districts and schools.
3. Monitoring of student test experience and the effects of new test administrations on students' school day and year.
4. Reporting of scores with more explanatory and framing information on what new scores signify.
5. Clarifying the use cases of different test designs for accountability versus instructional decisions.

OPPORTUNITY 1

Incorporating parent voices in assessment design and implementation decisions.

States have engaged in little outreach to parents when designing the new assessments. The limited outreach is, in part, because the assessments and their operational timelines were legislatively prescribed in Texas and Florida. Texas' House Bill 3906 required the TEA to develop its pilot program.⁵⁸ In Florida, Gov. DeSantis originally championed the proposal that required an operational progress monitoring assessment.⁵⁹ The intent of the resulting bill, Senate Bill 1048, was to prioritize measuring growth within the assessment system as well as having shorter tests, which limited the flexibility when Florida was designing the system.⁶⁰

With the exception of Texas, the states reported primarily receiving information and feedback from district officials on design decisions. In Texas, the state conducted focus groups with educators, parents, and students when planning the assessment's design. When recruiting for the focus groups, the TEA tried to account for location (i.e., urbanicity) and socioeconomic status, as well as ensuring representation from both traditional public schools and charter schools. The TEA also continues to collect information from students by administering optional short surveys to all students participating in the pilot directly after the students take the tests. Similarly, Texas educators and administrators provide feedback throughout the year in various ways, including surveys.

Delaware state officials do informally talk with students about the test design and the test experience as part of monitoring to ensure tests are being administered properly, but the state has not done specific outreach to parents. In speaking with Delaware parents, they believed that there is likely a system for parents to provide input but thought that any system would be difficult to navigate and, more important, the parents did not think that the state would use parent input.

The lack of outreach is important, because parents have differences in opinion on the usefulness of the test design. For example, although some Florida parents saw the value in having a within-year measure of student growth, other parents do not find all testing events valuable.

“I know that [opportunities to provide input] exist. So it’s almost like it’s a guarded secret, and you don’t like to talk about where those opportunities exist. And I think that even when parents do give input, I don’t think most of the parents that I’ve talked to in Delaware feel like they have a whole lot of faith that their opinions matter.”

—DELAWARE PARENT

OPPORTUNITY 2

Improving information flow between the state and parents, via districts and schools.

Parents cannot always keep up with information about state assessments, especially in recent years with multiple testing changes and disruptions due to the COVID-19 pandemic. Parents and community partners found it difficult to keep track of state assessments. Florida parents talked about how tests change every few years, to the point where they do not know which test is currently administered. A Delaware community partner noted the same issue with evolving assessments and feeling like she could not keep up with the changes.

The problem of switching assessments was compounded in Florida. A few parents were unaware that their children took state tests multiple times throughout the year. Others knew there was a new test but did not feel like it was much different than the previous test. Similarly, the Delaware parents we spoke with whose children participated in the social studies pilot testing were unaware of a new test.

Parents Want More Useful Information Beyond the Testing Calendar

Parents expressed an interest in better understanding the test but need additional resources. Many parents are not provided with information about the test and how the scores will be used. Florida parents reported that they are mostly told about the testing calendar and general tips related to testing, like getting a good night’s sleep or to eat breakfast before testing.

Although knowing when testing is occurring is important, parents in Florida and Delaware mentioned wanting to better understand what material was being tested. For instance, they wanted easier access to sample items. Although both states do have sample tests available, parents were not aware that they existed.

Similarly, both Florida and Delaware parents wanted more information about instructional resources so they could help their child at home. Some of the requests were for instructional materials before testing to help students review and prepare for the test and support their student learning. Parents were interested in instructional materials after testing to support student learning based on the student’s performance on the test.

Many of the parents we spoke with expressed frustration that they did not know how the scores were going to be used. This frustration was particularly prevalent in Florida, given the state’s third-grade literacy requirement that students must pass the grade 3 FAST or be retained. Parents of non-third-graders noted that they were concerned that scores would impact promotion and retention decisions.

Once parents receive their student’s score, they would benefit from additional resources to understand what the scores mean. The need for additional score reporting resources is particularly important as there are different expectations about student performance throughout the year. With the switch to through-year models where the assessment is not tied to a particular scope and sequence or curriculum, students are expected to grow in knowledge and skills throughout the year.



This is a switch from the end-of-year model, where the expectation is that students would score “proficient.” The state and districts are tasked with communicating these different expectations to parents to explain that it is no longer the expectation that students would score “proficient” in the fall or winter.

In some schools, Florida school staff explained to parents that there are different expectations for students across the three performance periods. For example, one Florida parent talked about attending an open house where the staff explained that students would not start the year at grade level, but by the end of the year, they would learn the skills to be either at or above grade level. Another parent described having a conversation with a vice principal after her child received a level 1 on the first test. The vice principal explained that the level 1 was where students start the year. In other Florida schools, there was little communication about how students should be progressing through the year.

“What happen[s] if a student got a level 1 or 2? What are the strategies the families can do to help the kids to move on to the next level?”

—FLORIDA PARENT

Updates to the Testing Program

These communication challenges are compounded when states make changes midyear while building and refining the pilots. Districts in Florida and Nebraska expressed frustration with midyear — and, in the case of Nebraska, year-to-year — changes, even if the changes were considered improvements by some.

In Florida, because of the short time to roll out the assessment, not all documents or policies were available

at the start of the year. The state also made changes to administration policies, particularly related to how much time students should test, based on feedback from the earlier testing events. One district said that the changes impacted students who needed accommodations the most, as there were changes to the extended time policy.

In Nebraska, the state announced changes to whether the fall or winter test could be used for proficiency determinations midway through the school year.

In addition to accountability consequences, this also impacted how certain districts were planning to use the data to identify students who no longer needed intervention services so that they could enroll in elective courses instead.

Similarly, Nebraska had changed the constraints on the computer adaptive testing engine, which selects items for students based on student responses to earlier materials to allow for items further from the student's enrolled grade level.⁶¹

One district assessment director in Nebraska noted that the changes have been particularly challenging for school districts already stretched for capacity, especially post-pandemic with staffing shortages and competing priorities. These changes, although necessary, create challenges for districts not only to implement the tests but also to communicate information about them to parents.

“As you can imagine, as a district leader we have many initiatives and don't always have the capacity and time to really study all the inner workings of the test. So that's been a challenge.”

—NEBRASKA DISTRICT LEADER

OPPORTUNITY 3

Monitoring of student test experience and the effects of new test administrations on students' school day and year.

The introduction of three, shorter tests a year has the potential to reduce student testing time, but so far, parents are not seeing those benefits.

Test Preparation

One reason the benefits of shorter testing time are not being realized is test preparation. Although there may be some benefits to test preparation (e.g., to familiarize students with the testing platform and types of questions they may see or to allow students with disabilities an opportunity to have guided reviews of their test accommodations), the increase in test events may also result in increased test preparation.⁶²

As districts have not provided parents with much information on testing, many parents were unsure if test preparation had increased. In Florida, some parents noted that there was test preparation before test events and assumed that it was before each of the tests, which would increase the total amount of test preparation over the course of a year. Delaware and Nebraska parents were unaware whether there was test preparation or not.

It should be noted that some parents were in favor of increased test preparation. There were Florida parents who thought that the amount and type of test prep was helpful for students to learn the grade-level material and become familiar with the testing platform. The parents who were more comfortable with the test preparation were also those whose child's school offered Saturday tutoring prior to the test based on students' prior test scores.

Changes to the School Schedule

In addition to test preparation interfering with instructional time, parents of students with disabilities were concerned about the impact of testing on the instructional day. Florida's FAST tests are untimed, and the first two administrations are designed to take 80 to 100 minutes. According to the state, the vast majority of the students completed the first test within 90 minutes; however, other students needed 2 to 2 1/2-hours and may continue testing throughout the day and into the following day. Schools have some flexibility in how they schedule students during the testing, particularly when some students have already completed the assessment. A Florida parent mentioned keeping her student with a disability home during the spring test because on days when other students were testing, he was asked to sit and watch movies all afternoon so as to not disturb testing that was in progress. Similarly, a Nebraska parent discussed how the increased number of test events interfered with services for her child who has a disability. The services were disrupted for a number of reasons, including that the student was in testing or that the student's providers were asked to proctor testing for other students. Parents of students with disabilities in both Florida and Nebraska said general changes to the school schedule were disruptive to their children, who need a predictable routine.

Test Day

The student experience on test day is also impacted by the state's decision whether or not to align the test with a particular scope and sequence of the standards.

In talking with students, a state official in Delaware noted that they appreciate being assessed right after instruction. Delaware parents were similarly supportive of having testing closer in time to instruction, particularly for social studies, which parents perceived as requiring more memorization than other subjects. Parents and educators also highlighted the potential benefits for students who move within the state during the school year. By having more consistent content within the state, the student may be less likely to miss instruction.

For the states opting to align with the state's standards, the first assessment of the year can be stressful for students, since they are testing on content they have not been taught. Anecdotally, the state has heard that high achievers find it stressful, as they are used to knowing answers to test questions. Florida parents used words like "discouraging," feeling "dumb," and "nerve-racking" to describe how their child felt after the first test. Nebraska parents were less familiar with the testing and did not have opinions on the relationship to when the material is taught and testing experience.

DOES IT FEEL LIKE MORE TESTING?

Delaware: Parents said that with the additional testing periods, it does feel like more testing even though the testing is shorter.

Florida: District leaders reported that it felt like more testing. Similarly, in a survey of nearly 400 Florida parents, 55% reported that the amount of testing felt the same as the previous year and 30% felt like testing time had increased.⁶³

Nebraska: Many of the Nebraska schools that opt to administer the fall and/or winter NSCAS Growth test continue to administer a commercially available interim assessment, which increases the total testing time.

OPPORTUNITY 4

Reporting of scores with more explanatory and framing information on what new scores signify.

A strength (and goal) of the new assessments is a faster turnaround for results. In Delaware, students receive results immediately after submitting their test. In Florida, families can access test scores through a parent portal within a few days after the student tests.

Despite the faster turnaround time, parents are not always receiving the results. Particularly when students are participating in pilot programs, parents may not receive their student's scores. In some cases, like the Delaware fourth- and sixth-grade tests, the reporting platform provides teachers with a raw score immediately after the student submits, but the pilot does not include a score release to families. The state may also be embargoing certain scores so that district leaders are unclear about which scores may or may not be shared with parents. In other cases, the lack of score-sharing with parents appears to be a strategic decision by the district. For instance, in Texas, districts are still learning how to use the scores and what the data means. In one school we spoke with, the principal noted that the school is focusing on teachers understanding the data first before they will consider sharing scores with families.

Some districts do communicate to parents that scores are available. For instance, in some Florida districts, the school sends scores home or sends parents text messages notifying them that scores are available. However, this practice is not universal. One Florida parent reported receiving her child's scores during an individualized education program meeting but not otherwise knowing they were available.

Despite the availability of scores, parents are not necessarily checking them. A Nebraska parent noted that she had not received her child's scores but likely could. She had not reached out to the school to receive the scores for test events from earlier in the year because those scores are not being used for any

instructional purpose. She noted that there are already so many things that she needs to keep track of — like grades, missing assignments, and field trip permission slips — and that NSCAS Growth scores are not a priority for her.

When parents do receive the results, trying to interpret them can be more challenging than with a typical end-of-year assessment. In Florida, the biggest challenge in interpreting results is understanding the results from the first test event in the fall where nearly all students receive a level 1, which is the lowest performance level. The state added percentile ranks and is working to include projected growth data on the score report to help educators and parents better understand how the student scored. However, Florida parents noted initial confusion over what to do with the information.

When confused about their child's score report, parents sometimes, but not always, turn to their child's teacher or school for additional information. Some parents wished their student's teacher would go over the test scores with parents to explain the scores or provide additional resources. Other parents were reluctant to require teachers to have one-on-one conversations with each family. For example, a parent of an eighth-grader noted that it would be unreasonable to ask teachers to meet individually with parents but thought that a parent group explaining the scores could be beneficial. To help support teachers in those conversations, states may develop additional materials. For instance, in Nebraska, an education services unit, which supports multiple districts, produced a video explaining the NSCAS Growth test and how to interpret the scores. The video is available for teachers to share with parents either outside of or during parent-teacher conferences.

A final challenge is interpreting the results in combination with other assessments. In Florida, many districts administer another commercially available interim assessment, and districts are discovering that the results of the interim assessment are not consistent with the results of the FAST. Nebraska faces a similar problem with the interim tool widely used within the

state. The discrepancy between the results of the interim test and the NSCAS Growth was one of the reasons Nebraska postponed the full implementation of NSCAS Growth, in which districts would be required to participate in all three test dates.

OPPORTUNITY 5

Clarifying the use cases of different test designs for accountability versus instructional decisions.

All through-year models have the potential to be beneficial over traditional end-of-year tests in terms of instructional relevance. They allow educators and parents to see student progress according to state standards earlier in the year and allow for reteaching. They may also provide the opportunity to reach proficiency earlier in the year, which may motivate students to take the earlier assessments more seriously and might eventually enable new forms of accelerated or competency-based instruction.

However, test design decisions may mean that some through-year models are more useful for instructional purposes than others.

Scope and Sequence

For the scope and sequence model, the test results allow educators to monitor students' learning based on what was taught. One Delaware teacher described the ability to provide more timely remediation in social studies. In talking with parents, they had not noticed that additional testing produced any changes in instruction. To these parents, the amount of instruction seemed the same as it was before the switch to the new testing program. However, given the assessment is still new, this may change in the future.

Delaware parents generally appreciated the idea of having the assessment administered soon after the content was taught. They did express some concern that having limited time periods for teaching the content may be challenging for certain classrooms.



Two parents described substantial teacher turnover in their children’s classrooms during the school year that was disruptive to learning and required additional flexibility within the school year. Another parent of a child with a disability noted that depending on the needs of the students in the classroom, they may need extra time to learn certain content.

Standards-Aligned With Only On-Grade-Level Items

The on-grade-only model, particularly early in the year, does not provide much actionable information. In Florida, most students receive the lowest performance level during the first test administration because they have not been taught the content. As noted earlier, the state is providing supplemental information such as percentile ranks and, in future years, projected growth, which could be useful to identify students who will need additional learning support to meet the end-of-year standards. Despite having the earlier test events, many Florida districts are continuing to use a commercially available formative assessment tool, which includes an instructional component, to provide more instructionally actionable information.

For the on-grade-only model, the second test event may be more helpful to inform instruction and the need for remediation. Florida parents described programs where their students received Saturday tutoring in the weeks before the spring test, with the tutoring based, in part, on their PM2 scores. However, not all Florida parents felt like schools were doing much with the scores. One Florida parent said, “Once you see those results, what’s being done about it? For example, my younger son, he’s showing a decline in progression. And it’s like, okay, great. Now what do I do with this? Like, what does this mean? I don’t know.”

Another parent whose child did not pass the third-grade test said, “If they are going to continue with that test, there should be tutoring, and they should pay for it. The state should pay for additional tutoring or something because all children don’t learn at the same pace. Since the pandemic, things have changed drastically for kids and adults.”

Standards-Aligned With Some Off-Grade-Level Items

The off-grade-level model can potentially provide additional instructional information. Nebraska decreased the restraints on the computer-adaptive testing engine to allow students to receive items farther from grade level on the earlier assessments. The state made the change to provide more instructionally actionable information.

However, because of the way the scores were initially reported, the scores had limited usefulness for instruction. Some of the reporting issues were related to the reporting platform itself, which did not allow users to organize the data in actionable ways. Other reporting issues are related to the types of contextual information the state reports. For instance, the state did not have the normative growth data that districts were accustomed to receiving from commercially available interim assessments. As of spring 2023, the state has updated the reports, but reporting initially was a source of challenge.

There also may be limited utility because of the small number of items on the assessment. The test is too short to be useful for either purpose. Districts we talked with continue to rely on a commercially available interim test for student goal-setting and instruction.

All through-year models have the potential to be beneficial over traditional end-of-year tests in terms of instructional relevance. However, test design decisions may mean that some through-year models are more useful for instructional purposes than others.

Trade-Offs Between Instruction and Accountability

The trade-off for using on- versus off-grade-level items is the ability to have the earlier tests used for accountability. Florida and Texas both considered including off-grade items but did not because they wanted to use the earlier testing events for federal accountability purposes and to better measure student growth based on grade-level standards. Nebraska tried to integrate both — having instructional and accountability information — but ultimately opted for instructional purposes for the earlier tests. Nebraska officials made the change because of the challenges in meeting the on-grade-level blueprint and because

they wanted to have items that could better help inform instruction.

States using a standards-aligned approach are finding that they need to choose between using the earlier test events for instruction or for accountability. As one Nebraska district leader said, “[We] can’t get the best things of both when the purposes are different.”

Similarly, a Florida community member noted, “[Regardless of the model] if the inputs don’t change, the scores and gaps won’t change either.” If educators are not using the information from the tests to adjust instruction, there will not be improvements in learning.

CHALLENGING LOGISTICS OF TEST ADMINISTRATION

The switch to the more frequent tests that are used as accountability tools can make the logistics of administration more challenging in states and districts across the country.

Nebraska districts said that with the interim assessment the school administers, a person walking into the school would not notice that the assessment was underway. However, with the state assessment, the test becomes a disruption — not only to the instructional day but also to staff needing to prepare to administer the test.

When the assessment is being used for accountability, administrative issues such as registering students for the test (also called “rostering”) become more cumbersome as districts must ensure that a student is registered only once for the test. When there is student mobility among schools or districts, the district must update its rostering to ensure the student is associated with the correct school and district. For the commercial interim assessment, the assignment registration is less complicated because the contract is with a particular district. If a student is registered for the test in multiple districts, they complete it and receive the score report only in the district where they tested. There are no concerns that the student did not participate in testing, as there would be for the test used for federal accountability purposes. Because the test is for accountability, Nebraska requires additional test security measures such as removing posters from walls and rearranging desks in the classroom, which increases teacher time to prepare for the test.

In Florida, the issue is that multiple test vendors are responsible for different grade levels of the assessment system. One vendor provides the assessment for grades 3-8, which is used for federal accountability. Separate vendors provide the assessments for earlier grade levels in pre-K through grade 2. Because there are multiple vendors, there are also multiple testing platforms, which makes it more challenging to administer.

Why Not Leave Interim Assessments to School Districts?

Many districts administer their own interim assessments. Interim assessments are administered periodically during the year and, depending on the assessment, can serve a formative function (i.e., for learning) or a summative function (i.e., to measure how much a student has learned).

Districts administer interim assessments to complement the state assessment(s). The district-level assessments cannot be used as part of the federal accountability system without meeting ESSA assessment requirements, such as all students within the state taking the same assessment and the alignment with state standards. Additionally, many states leave these decisions to districts because districts have primary purview over curriculum adoption and the scope, sequencing, and measurement of content during the school year.

With state-designed through-year assessments serving as an interim assessment, there may be some improvements to student testing experiences.

One area where state involvement may improve student experience is accommodations. State-developed interim assessments have the potential to increase the alignment across the assessments. With commercially developed interim assessments, there may be a different set of accommodations offered between the commercially developed and state-created assessment. Better alignment leads to more opportunities for students to practice with the accommodations. Students not only have more opportunities to practice with the accommodations they will be testing with, but there is also the potential for better alignment between instructional and assessment accommodations because the accommodations are being used more frequently. There is also the potential for more robust accommodations than with commercially available interims. For example, Florida plans to provide American Sign Language accommodation and print materials in future years, which are not typically available on commercial interim assessments.

State-developed interims may also have greater alignment with the state's specific content standards. Because the assessments are state-developed, the

items can better target the state's standards. This is more relevant in instances where the state has unique standards that may not be represented in the test vendor's item bank.⁶⁴

Having a single state-adopted interim assessment could have other benefits. It provides a standardized metric across the state, allowing for better monitoring and identifying students and schools for support. There is also the opportunity for additional resources for districts. For instance, by creating the assessment, states have the opportunity to offer districts more resources and opportunities to provide technical assistance in how to use interim assessment data appropriately, both in terms of what the data can and cannot say, as well as how to use it for instruction.

Simply adopting a state interim assessment does not mean that districts will not continue to use their own interim assessment. As we have seen in the states adopting through-year assessments, districts continue to administer their own interim assessment. This is partly due to familiarity with the assessments and the data. Another reason is that districts appreciate having longitudinal data to track student progress. If the state opts for a different assessment, districts lose that historical data.

Recommendations

Conversations with state and district leaders and parents in Delaware, Florida, Nebraska, and Texas highlight the following recommendations for state and federal policymakers looking to adopt or improve through-year assessment systems, and many of the recommendations are also applicable to state assessment systems more generally.

1. Design new assessment systems to solve an agreed-upon problem identified by districts and their communities.

As state leaders consider adopting a through-year assessment model, it is important that the test solve an agreed-upon problem.

Nebraska's state-led initiative, though in response to district concerns about testing time and speed of reporting, has seen substantial delays in transitioning from a pilot to a statewide model with full buy-in. Part of the delay is due to hesitancy by districts because the state assessment was slow to meet or surpass the quality of data and reporting the districts were used to with the commercially available assessment. Similarly, in Florida and Texas, the initiatives were driven by state legislation without a clear local push.

In Delaware, the state was clear that the work of the Social Studies Coalition was instrumental in making the change. There was a known problem with the grade-span approach to social studies that the new assessment was designed to address.

Despite the success in social studies, Delaware is not considering making the same switch with the state's English language arts and math assessment. Part of the reluctance is that the test vendor already provides an optional interim assessment, but more important, districts are not advocating for the change as they were for social studies.

Before shifting to a through-year assessment system, state policymakers should determine whether the through-year model will address agreed-upon problems or whether the shift will create new problems. It is also important for policymakers to approach the test design process by first considering the test experiences and unmet needs of underserved students and families when designing the assessment (e.g., considering how changes to schedules impact routines and services for students with disabilities or how students will be spending their day when not engaged in testing). By thinking through underserved student experiences earlier, policymakers can make design decisions to better serve all students and anticipate points of friction that may detract from the overall goals of the new assessment.

2. Increase family and community engagement at multiple steps in the assessment design process.

Although Texas held focus groups to better understand the competing demands for test score usage, the other states analyzed did little visible parent or community engagement, despite one reason for changing assessments was parent dissatisfaction with testing time or student experience under prior models. Because of the lack of parent engagement, most parents we spoke with felt like they did not have a voice when it came to assessments, and they would like to have one.



Stronger communications could better tailor the assessment to the needs of the community. For example, if the community is concerned about within-year growth, the state may consider a standards-aligned model similar to Florida. If communities are more concerned about the consistency of when content is taught across the state, the state may consider a scope and sequence model similar to Delaware. It could also help avoid misunderstandings or distrust of the assessment and the scores. For instance, parents highlighted not understanding the purpose of testing and expressed concerns that they were unsure how the scores would be used.

Outreach should not only be limited to selecting the model. It should include other aspects of test design, particularly the score report design process. By engaging parents and the community when designing score reports, the state can develop clearer score reports and identify supplemental materials to support parents in understanding their child's progress. This is particularly important for through-year assessments, since the meaning of student performance and progress changes over the course of the year. These conversations may also uncover additional areas where parents would like more information, such as clearer information about what is being tested.

Given the need for stronger communications, state policymakers should be more intentional in engaging families and communities throughout the assessment design process. In engaging with parents and community members, states should implement outreach strategies that include multiple avenues for engagement. In addition to focus groups, states should consider activities such as presentations to community groups, written communications, surveys, and listening tours. This process of deeper engagement will require a critical investment of time and resources.

3. Provide districts with timely, robust resources to implement and disseminate information about new tests.

Districts largely rely on state resources for communication. They need timely, robust resources to ensure correct implementation of the testing program and communicate changes in test administration, uses, and scoring with their teachers, parents, and communities.

Florida is still building out resources as tests are being administered, which has caused some challenges in determining what information can be shared (or not shared) with families and communities. Nebraska districts faced a similar problem in understanding what information can be shared and when. And some pilot districts in Texas have not shared results with families while tests are still new and not well understood.

Given the reliance on state communications, state leaders should prioritize creating timely, easily understandable resources for districts and families. States should also provide technical assistance to districts, including opportunities to engage in a community of practice with other districts grappling with similar questions.

4. Implement formal continuous improvement processes based on monitoring of student experiences.

ESSA requires that states monitor testing to ensure proper implementation of the tests;⁶⁵ however, states will not monitor every school each year. States largely rely on districts to provide feedback on student experiences during testing. In the districts we spoke with, few have formal processes to monitor student experience during testing, and they rely on reports from principals and school test coordinators. Given that through-year assessments are new, states and districts should make a concerted effort to identify unintended consequences related to the tests by more thoroughly monitoring and evaluating testing. For example, states could select a larger sample of schools and districts to monitor in each test administration. This monitoring should focus not only on test security and administration fidelity, but also on student experiences, with an emphasis on test experiences among students needing accommodations or historically lower-performing student groups.

State and district leaders should implement these formal monitoring processes to better understand the impact of assessments on students and then make ongoing improvements to the assessment system based on those findings.

State and district leaders should implement these formal monitoring processes to better understand the impact of assessments on students and then make ongoing improvements to the assessment system based on those findings.

5. Explore making changes to federal law to achieve the stated goals of through-year assessments.

States are adopting through-year assessments with the goals of:

- Reducing testing time.
- Creating more timely, instructionally useful results.
- Improving accountability (where results can also be used for proficiency and new growth measures).

These stated goals magnify the tension between the accountability and instructional uses of the test scores. When trying to promote accountability uses, the tests become less instructionally relevant unless the state can require a particular scope and sequence within a school year. Because most states leave decisions about scope and sequence and curriculum to the individual districts, they will continue to face these tensions.

Some of the tension is codified in ESSA. For instance, ESSA requires that state summative tests provide diagnostic information that can be used for instructional purposes.⁶⁶ If states were only required to have a score used for accountability, the state might be able to shorten the end-of-year test and reallocate resources to build out additional assessments for instructional purposes. Similarly, ESSA requires that scores only include on-grade-level material, which limits a state's ability to both include more targeted information to guide instruction and to have a shorter test.⁶⁷ Given these tensions, experts have been identifying additional opportunities outside of through-year assessments to provide creative solutions that can still produce an accountability score.⁶⁸

Federal policymakers should consider the trade-offs of different approaches to state assessment when thinking about ESSA reauthorization.

Conclusion

Creative solutions and new models are necessary in the world of state assessment and accountability to address legitimate criticisms and shortcomings, while continuing to support the important priorities and purposes served by state assessment results. State interest in through-year assessments acknowledges that the end-of-year model can provide valuable information but cannot provide all of the information stakeholders want or need.

Given demonstrated and growing state interest in exploring and implementing through-year assessment models, this report provides additional guidance to state decision-makers, as well as test vendors and others who are interested in the implications of state test design and policy change. That said, this report is only a start. Additional research is needed, particularly with ELs and their families and highly mobile students, as well as with teachers, especially teachers of color and those in underserved communities. Further, as many of the testing programs we studied are still in pilot stages, similar research should be conducted once the assessments are operational to monitor for unintended consequences and evaluate the implementation of state through-year assessments.

Beyond considerations for through-year assessments alone, this report raises broader questions about the specific limitations and opportunities available for assessment innovation within current federal law. It offers lessons for state leaders considering any kind of major education policy shift, assessment-related or otherwise, to consider, engage, and empower families and communities who are most directly impacted by the change and most often left out of decision-making. ✦

Appendix

PARTICIPATING DISTRICTS AND COMMUNITY GROUPS

In addition to talking to district leaders in each state, the following organizations participated in this analysis from March 2023 to July 2023, either through direct interviews or through recruiting parents to be interviewed (Disclosure).

Note: *Because Texas was in the early stages of its pilot, we spoke exclusively with district leaders.*

National

- Easterseals (via state affiliates)
- National Parents Union

Delaware

- Five participating parents
- DelawareCAN
- DELLTA
- EastSide Charter School
- Las Americas ASPIRA Academy
- Network Connect
- Rodel Foundation
- Social Studies Coalition of Delaware
- United Way of Delaware

Florida

- 12 participating parents
- The Children’s Trust
- ExcelinEd
- Helios Education Foundation
- Lastinger Center for Learning at the University of Florida
- Overtown Youth
- South Brevard NAACP

Nebraska

- Three participating parents
- PTI Nebraska

Endnotes

- 1 Dave Powell, Senna Lamba, Khaled Ismail, and Joshua Marland, *What Are Through-Year Assessments? Exploring Multiple Approaches to Through-Year Design*, Education First, 6, <https://www.education-first.com/library/publication/what-are-through-year-assessments/>.
- 2 Improving the Academic Achievement of the Disadvantaged, 200 U.S.C. § 6301 et seq., <https://www.law.cornell.edu/uscode/text/20/chapter-70/subchapter-1>.
- 3 ESSA, 20 U.S.C. § 6311 (b)(2)(B), <https://www.law.cornell.edu/uscode/text/20/6311>.
- 4 Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*
- 5 Through-year assessments also have the potential to have a greater impact on English learners and highly mobile students. Although we specifically asked state and district officials about these groups, additional research is needed to better understand how parents and students experience the new assessments.
- 6 Michelle Croft, Haily T.N. Korman, and Titilayo Tinubu Ali, “Demystifying Statewide Standardized Assessments,” Bellwether, 2023, <https://bellwether.org/publications/demystifying-standardized-assessments/>.
- 7 “What Parents Need to Know About Standardized Tests,” EdNavigator, <https://www.ednavigator.org/resources/what-parents-need-to-know-about-standardized-tests>; Michelle Croft, “Confused by Your Child’s State Assessment Results? You’re Not Alone.,” Bellwether, November 10, 2021, <https://bellwether.org/opinion/confused-by-your-childs-state-assessment-results-youre-not-alone/>.
- 8 Jason Weeby, “Creating More Effective, Efficient, and Equitable Education Policies With Human-Centered Design,” Bellwether, February 27, 2018, <https://bellwether.org/publications/creating-more-effective-efficient-and-equitable-education-policies-human-centered-design/>.
- 9 Stephen Kostyo, Jessica Cardichon, and Linda Darling-Hammond, *Making ESSA’s Equity Promise Real: State Strategies to Close the Opportunity Gap*, Learning Policy Institute, 2018, <https://learningpolicyinstitute.org/product/essa-equity-promise-report>.
- 10 ESSA, 20 U.S.C. § 6311 (b)(2)(B), <https://www.law.cornell.edu/uscode/text/20/6311>.
- 11 ESSA, 20 U.S.C. § 6311 (b)(2)(B), <https://www.law.cornell.edu/uscode/text/20/6311>; Croft, Korman, and Tinubu Ali, “Demystifying Statewide Standardized Assessments.”
- 12 ESSA, 20 U.S.C. § 6311 (c), <https://www.law.cornell.edu/uscode/text/20/6311>.
- 13 Abigail Swisher and Patricia Saenz-Armstrong, “State of the States 2022: Teacher and Principal Evaluation Policies,” National Council on Teacher Quality, 2022, <https://www.nctq.org/publications/State-of-the-States-2022:-Teacher-and-Principal-Evaluation-Policies#ack>.
- 14 Daarel Burnette II, “When States and the Feds Take Different Paths on School Accountability,” *Education Week*, June 12, 2018, <https://www.edweek.org/policy-politics/when-states-and-the-feds-take-different-paths-on-school-accountability/2018/06>.
- 15 “Principles and Standards for Quality Charter School Authorizing,” National Association of Charter School Authorizers, 2023, https://qualitycharters.org/wp-content/uploads/2023/03/Principles-and-Standards_2023.pdf.
- 16 For example, Minnesota describes how school leaders can use the information to identify professional development needs and educators can use the information to identify strengths or gaps in curriculum as an intended use of the state’s MCAS test. “Use Statewide Assessment Data,” Minnesota Department of Education, n.d., <https://testing123.education.mn.gov/test/analyze/report/>.
- 17 Laura Hamilton, Richard Halverson, Sharnell S. Jackson, Ellen Mandinach, Jonathan A. Supovitz, and Jeffrey C. Wayman, *Using Student Achievement Data to Support Instructional Decision Making*, IES Practice Guide, 2009, https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm_pg_092909.pdf.
- 18 Randy E. Bennett, “Opt Out: An Examination of Issues,” *ETS Research Report* 2016, no. 1 (2016): 1–16, <https://onlinelibrary.wiley.com/doi/full/10.1002/ets2.12101>.
- 19 Andrew Ujifusa, “Amid Cries of Overtesting, a Crazy Quilt of State Responses,” *Education Week*, July 8, 2015, <https://www.edweek.org/policy-politics/amid-cries-of-overtesting-a-crazy-quilt-of-state-responses/2015/07>.
- 20 Katherine E. Castellano and Andrew D. Ho, *A Practitioner’s Guide to Growth Models*, Council of Chief State School Officers, 2013, https://scholar.harvard.edu/files/andrewho/files/a-practitioners_guide_to_growth_models.pdf.
- 21 Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*
- 22 Ibid.
- 23 ESSA, 20 U.S.C. § 6311 (b)(2)(B)(viii)(II), <https://www.law.cornell.edu/uscode/text/20/6311>.
- 24 For a full list of decisions ranging from the design to logistical considerations, see Nathan Dadey, Carla Evans, and Will Lorie, “Through-Year Assessment: Ten Key Considerations,” Center for Assessment, 2023, <https://www.nciea.org/library/through-year-assessment-ten-key-considerations/>.
- 25 Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*
- 26 ESSA, 20 U.S.C. § 6311 (b)(a)(A), <https://www.law.cornell.edu/uscode/text/20/6311>.
- 27 Dadey, Evans, and Lorrie note that many of these decisions are related to one another. The decision about connection to curriculum is closely related to the purpose of testing. Dadey, Evans, and Lorie, “Through-Year Assessment: Ten Key Considerations.” See also Nathan Dadey and Aneasha Badrinarayan, “In Search of the ‘Just Right’ Connection Between Curriculum and Assessment,” Center for Assessment, April 21, 2022, <https://www.nciea.org/blog/in-search-of-the-just-right-connection-between-curriculum-and-assessment/>; Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*

- 28 Kathleen Porter-Magee, "PARCC Eliminates Through-Course Assessments," Thomas B. Fordham Institute, June 30, 2011, <https://fordhaminstitute.org/national/commentary/parcc-eliminates-through-course-assessments>; Catherine Gewertz, "State Consortium Scales Back Common-Test Design," Education Week, July 7, 2011, <https://www.edweek.org/teaching-learning/state-consortium-scales-back-common-test-design/2011/07>.
- 29 Sarah Schwartz, "The Case for Curriculum: Why Some States Are Prioritizing It With COVID Relief Funds," *Education Week*, September 20, 2022, <https://www.edweek.org/teaching-learning/the-case-for-curriculum-why-some-states-are-prioritizing-it-with-covid-relief-funds/2022/09>; *Incentivizing Smart Choices: How State Procurement Policies Can Promote the Use of High-Quality Instructional Materials*, Chiefs for Change, April 2021, <https://www.chiefsforchange.org/wp-content/uploads/2021/04/Incentivizing-Smart-Choices-How-State-Procurement-Policies-Can-Promote-the-Use-of-High-Quality-Instructional-Materials.pdf>.
- 30 Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*
- 31 "Innovative Assessment Demonstration Authority," U.S. Department of Education, <https://www2.ed.gov/admins/lead/account/iada/index.html>.
- 32 Approximately 65% of the parents we spoke with were persons of color, and of those who disclosed their child's disability status, 69% were a parent of a student with a disability.
- 33 We learned through the recruitment process that in-person interviews facilitated with a Spanish speaker, in which participants are recruited through community organizations, may be more successful than interviews conducted remotely, but we were not able to fully implement those recommendations within our research timeline.
- 34 "Delaware Transitions to Innovative Through Assessment System for Social Studies by SY2024-2025," Delaware Department of Education, <https://www.doe.k12.de.us/cms/lib/DE01922744/Centricity/Domain/111/SS%20Through%20Assessment%20System.pdf>.
- 35 *Online Through Year Administration Manual for DeSSA Social Studies*, Delaware Department of Education, 2022, http://delaware.pearsonaccessnext.com/resources/resources-training/DE1143424_DESSA_TY22-23_G8SS_TAM_WebTag.pdf.
- 36 Ibid.
- 37 Ibid.
- 38 "Governor DeSantis Announces End of the High-Stakes FSA Testing to Become the First State in the Nation to Fully Transition to Progress Monitoring," Florida Office of the Governor, September 14, 2021, <https://www.flgov.com/2021/09/14/governor-desantis-announces-end-of-the-high-stakes-fsa-testing-to-become-the-first-state-in-the-nation-to-fully-transition-to-progress-monitoring/>.
- 39 Student Assessments, Florida Senate Bill 1048 (2022), <https://www.flsenate.gov/Session/Bill/2022/1048>.
- 40 "K-12 Student Assessment: FAST Assessments," Florida Department of Education, <https://www.fldoe.org/accountability/assessments/k-12-student-assessment/best/>.
- 41 "2022-23 Florida Assessment of Student Thinking (FAST) Fact Sheet: Grades 3-10 ELA Reading and Grades 3-8 Mathematics," Florida Department of Education, <https://www.fldoe.org/core/fileparse.php/20102/urlt/2223FASTGrd310FS.pdf>.
- 42 Ibid.
- 43 Ibid.
- 44 Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*
- 45 "NSCAS Growth Frequently Asked Questions," Nebraska Department of Education, July 2021, <https://www.education.ne.gov/wp-content/uploads/2021/07/NSCAS-Growth-FAQ-JULY21-Final2.pdf>.
- 46 "Nebraska's Transition to NSCAS Growth," Nebraska Department of Education, December 19, 2022, <https://www.education.ne.gov/wp-content/uploads/2022/12/NSCAS-Growth-FAQ-12-19-22.pdf>.
- 47 "State Board Report: January 2023," Nebraska Department of Education, <https://www.education.ne.gov/boardreport/january-2023/>.
- 48 "2023-2024 NSCAS Testing Windows," Nebraska Department of Education, July 21, 2023, https://www.education.ne.gov/wp-content/uploads/2023/07/2023-2024_NSCAS_Test_Window-7-21-23.pdf.
- 49 Joe Dejka and Lauren Wagner, "New Nebraska State Test Scores Show Extent of COVID Academic Slide," *Omaha World-Herald*, November 23, 2022, https://omaha.com/news/local/education/new-nebraska-state-test-scores-show-extent-of-covid-academic-slide/article_589e80b0-6ad3-11ed-8e21-e377538afe85.html.
- 50 "2022-2023 Schedule Guidance for NSCAS Growth and NSCAS Science," Nebraska Department of Education.
- 51 Ibid.
- 52 H.B. 3906, 86th Texas Legislature (2019), <https://legiscan.com/TX/text/HB3906/id/2027989>.
- 53 Powell, Lamba, Ismail, and Marland, *What Are Through-Year Assessments?*
- 54 "Texas Through-Year Assessment Pilot," Texas Education Agency, <https://tea.texas.gov/student-assessment/assessment-initiatives/hb-3906/texas-through-year-assessment-pilot>.
- 55 Ibid.
- 56 "2022-2023 Texas Through-Year Assessment Pilot (TTAP) FAQs," Texas Education Agency, updated September 16, 2022, <https://tea.texas.gov/student-assessment/testing/2022-ttap-faq.pdf>.
- 57 Ibid.
- 58 H.B. 3906, 86th Texas Legislature (2019), <https://legiscan.com/TX/text/HB3906/id/2027989>.
- 59 "Governor DeSantis Announces End of the High-Stakes FSA Testing to Become the First State in the Nation to Fully Transition to Progress Monitoring."
- 60 Student Assessments, Florida Senate Bill 1048 (2022), <https://www.flsenate.gov/Session/Bill/2022/1048>.
- 61 Croft, Korman, and Tinubu Ali, "Demystifying Statewide Standardized Assessments."

- 62 One district content specialist noted that for students with disabilities, having additional opportunities to practice with their accommodations can be helpful, particularly for the text-to-speech accommodation, so the student is more comfortable using the accommodation prior to testing.
- 63 As part of the interview registration process, we asked parents to complete a short screener survey to ensure they were eligible to participate in the study. Parents were also asked if they were aware of the new test and if they thought testing time had increased, decreased, or stayed the same. A total of 372 parents responded to the change in testing time question and racial identity questions. Approximately 57% of the parents who responded to the survey identified as a person of color.
- 64 See for example, Alabama’s Course of Study, which diverges from the Common Core of State Standards. “Alabama Course of Study English Language Arts,” Alabama Department of Education, 2021, <https://www.alabamaachieves.org/wp-content/uploads/2021/08/2021-Alabama-English-Language-Arts-Course-of-Study.pdf>.
- 65 “A State’s Guide to the U.S. Department of Education’s Assessment Peer Review Process,” U.S. Department of Education, September 24, 2018, <https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>.
- 66 Scott Marion and Derek Briggs, “Just Give Us a Little,” Center for Assessment, July 13, 2022, <https://www.nciea.org/blog/just-give-us-a-little/>.
- 67 ESSA, 20 U.S.C. § 6311 (b)(2)(B); “A State’s Guide to the U.S. Department of Education’s Assessment Peer Review Process.”
- 68 Marianne Perie, “Is the School Assessment and Accountability Era Over? Insights From Marianne Perie, Director of Assessment Research and Innovation at WestEd,” WestEd, March 23, 2023, <https://www.wested.org/wested-bulletin/insights-impact/is-the-school-assessment-and-accountability-era-over-marianne-perie/>.

About the Authors



MICHELLE CROFT

Michelle Croft is an associate partner at Bellwether in the Policy and Evaluation practice area. She can be reached at michelle.croft@bellwether.org.



TITILAYO TINUBU ALI

Titilayo Tinubu Ali is a partner at Bellwether in the Policy and Evaluation practice area. She can be reached at titilayo.ali@bellwether.org.



BONNIE O'KEEFE

Bonnie O'Keefe is a senior associate partner at Bellwether in the Policy and Evaluation practice area. She can be reached at bonnie.okeefe@bellwether.org.

About Bellwether

Bellwether is a national nonprofit that exists to transform education to ensure systemically marginalized young people achieve outcomes that lead to fulfilling lives and flourishing communities. Founded in 2010, we work hand in hand with education leaders and organizations to accelerate their impact, inform and influence policy and program design, and share what we learn along the way. For more, visit bellwether.org.

DISCLOSURE

Bellwether works with organizations and leaders who share our viewpoint—diverse commitment to improving education and advancing equity for all young people — regardless of identity, circumstance, or background. As part of our commitment to transparency, a list of Bellwether clients and funders since our founding in 2010 is publicly available on our website. An organization's name appearing on our list of clients and funders does not imply any endorsement of or by Bellwether.

In the context of this report, Bellwether has collaborated with other organizations mentioned in various capacities not necessarily focused on through-year assessments.

ACKNOWLEDGMENTS

We would like to thank the many individuals who gave their time and shared their knowledge with us to inform our work, including EduDream and its funders, the Bill & Melinda Gates Foundation and the Walton Family Foundation, for their financial support of this project. We also want to thank the EduDream advisory group, especially Gloria Corral and Carlos Robledo at Parent Institute for Quality Education, for their additional support as well as the EduDream grantees cohort.

This report is part of a research series on innovative assessment approaches and alternative accountability models coordinated under The K12 Research for Equity Hub (The Hub). The Hub is managed by EduDream and funded by the Bill & Melinda Gates Foundation and the Walton Family Foundation. No personnel from the Bill & Melinda Gates Foundation or the Walton Family Foundation participated in the creation of The Hub's research.

We would also like to thank our Bellwether colleagues Thomas Gold, Marisa Mission, and Alexis Richardson for their input and support. Thank you to Kate Neifeld, Andy Jacob, Zoe Campbell, Julie Nguyen, and Amber Walker for shepherding and disseminating this work, and to Super Copy Editors.

The findings and conclusions contained in this report are those of the authors and do not necessarily reflect the positions and/or policies of the Bill & Melinda Gates Foundation or the Walton Family Foundation. The contributions of these individuals and entities significantly enhanced our work; however, any errors in fact or analysis remain the responsibility of the authors.



© 2023 Bellwether

- Ⓒ This report carries a Creative Commons license, which permits noncommercial re-use of content when proper attribution is provided. This means you are free to copy, display and distribute this work, or include content from this report in derivative works, under the following conditions:
- ① **Attribution.** You must clearly attribute the work to Bellwether and provide a link back to the publication at www.bellwether.org.
- Ⓓ **Noncommercial.** You may not use this work for commercial purposes without explicit prior permission from Bellwether.
- Ⓒ **Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

For the full legal code of this Creative Commons license, please visit www.creativecommons.org. If you have any questions about citing or reusing Bellwether content, please contact us.