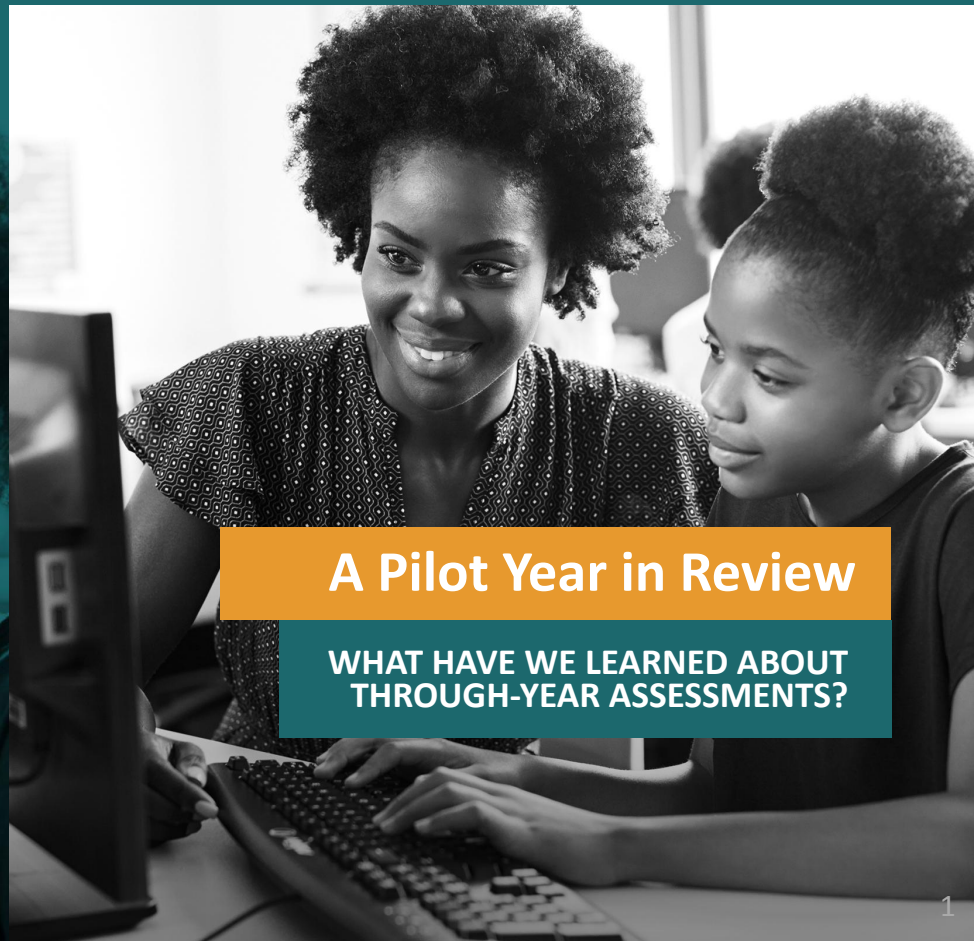
Innovations *in* Assessment

educationfirst

BILL & MELINDA GATES *foundation*

WALTON FAMILY FOUNDATION



A Pilot Year in Review

WHAT HAVE WE LEARNED ABOUT THROUGH-YEAR ASSESSMENTS?

About this publication

A Pilot Year in Review: What have we learned about through-year assessments?

This publication interrogates key takeaways from through-year assessment pilots administered during the 2022-2023 school year. We explore key design decisions, enabling conditions and implications for future research and practice. This publication is part of a series published through Education First's Through-year Curriculum-Connected Assessment Grant Program.

Authors

Dave Powell
Senior Consultant
Education First

Emma Fortier
Associate
Education First

Senna Lamba
Associate
Education First

Khaled Ismail
Principal
Education First

Directory

Executive Summary

Why Through-Year?

Lessons Learned

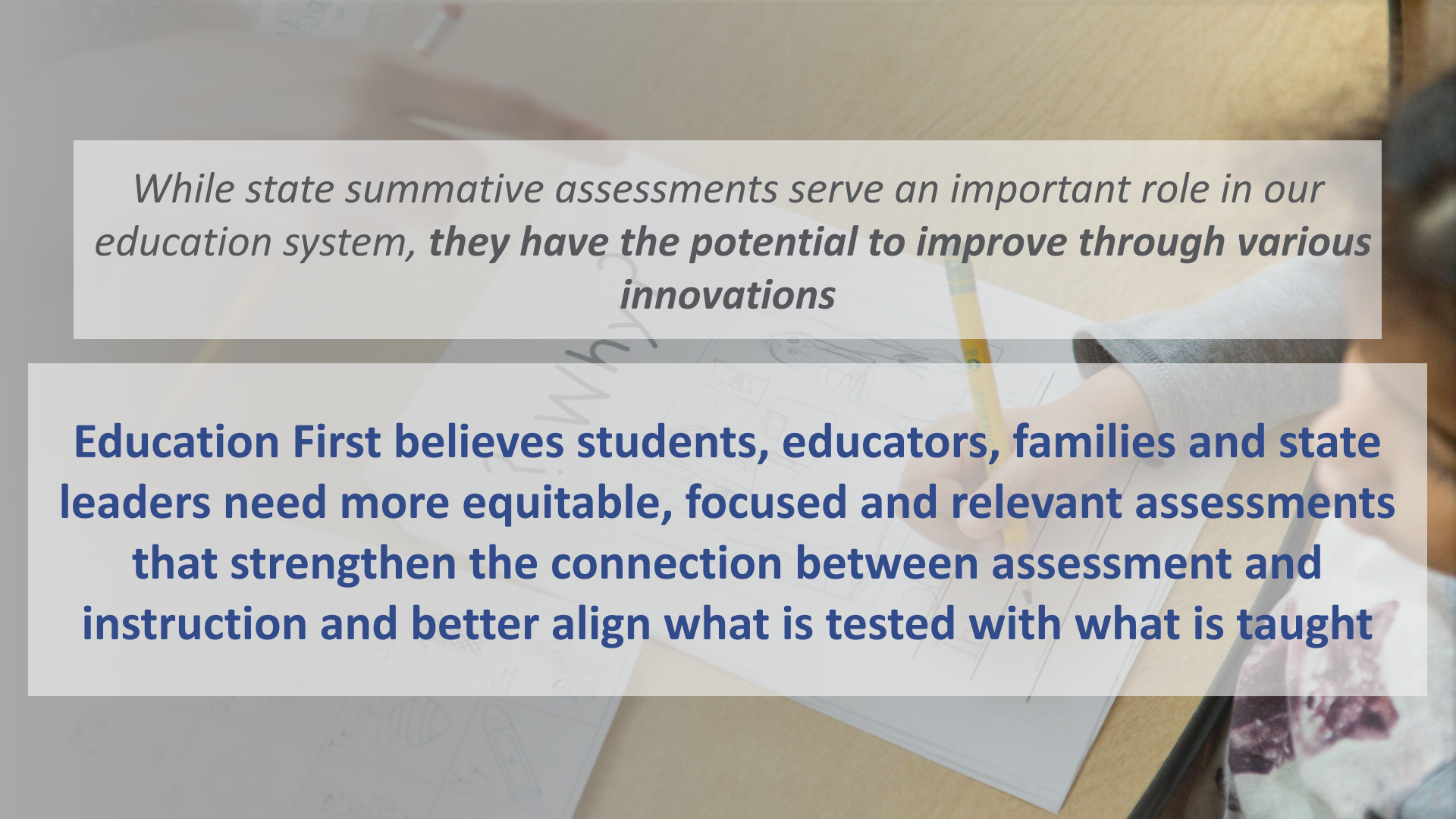
Full Report

Exploring the Models

Return to the landing page

Table of Contents

Click on these titles to navigate to other sections within this presentation.



While state summative assessments serve an important role in our education system, ***they have the potential to improve through various innovations***

Education First believes students, educators, families and state leaders need more equitable, focused and relevant assessments that strengthen the connection between assessment and instruction and better align what is tested with what is taught

Many states are exploring through-year assessments to address some long-standing, legitimate concerns about traditional end-of-year summative assessments

Stakeholders (including students, families and educators) often see traditional end-of-year summative assessments as:

- **Lacking utility to teaching and learning**
- **Providing untimely results that do not inform instruction**
- **Requiring a large footprint on the overall system**
  (in terms of the resources needed, time for preparation and administration)
- **Misaligned to what and when students are taught and their curriculum**

Read more about the reasons for the growing interest in through-year assessment models in Education First's publication, "What are Through-year Assessments?"

Sources: Education First (2022). Marion, S. (2021).

In this report, we use the following definitions that build on our prior thinking and the work of others in this field:

Through-Year Assessment Models	Through-year assessment models administer multiple tests throughout the school year as part of an assessment system designed to produce a single summative score meeting federal and state accountability requirements. Through-year assessment models are also referred to as "through-course" by some states.
Curriculum-Agnostic Approach	Through-year assessment models that **test the entire content domain (or grade-level standards) throughout the year at each testing administration**, and do not try to align content tested to curriculum.
Curriculum-Aligned Approach*	Through-year assessment models that directly **draw on the content found in specific curriculum**. This model is also referred to as "curriculum-specific" or "curriculum-embedded."
Curriculum-Relevant Approach	Through-year assessment models that can be **flexibly aligned with multiple curricula, a scope and sequence or pacing of content**. This approach is also referred to as "scope and sequence aligned", "instructionally relevant" or "instructionally aligned."

*In previous publications, we referred to this as curriculum-specific embedded

Sources: Education First (2022) Dadey and Badrinarayan (2022) Dadey, N., & Gong, B. (2023)

With these definitions in mind, this publication explores the following research questions

Primary Research Questions:

What are the lessons learned from a group of states who piloted through-year models in the 2022-2023 school year?

What are key implications and recommendations for scaling the models?

What are the outstanding questions, needs and considerations for the future of innovations in assessment?

Across the states we reviewed for this publication,

over

2.5 million

students tested using a through-year assessment during the 2022-2023 school year.

WHY THROUGH-YEAR ASSESSMENT MODELS?

THE POTENTIAL BENEFITS, TENSIONS AND ONGOING ISSUES

The Research Supporting Through-Year Assessments

Education First believes that if states test throughout the year, align those tests to what is taught and provide timely reporting–student learning will improve

If you…

- Test throughout the year
- Align the content of the test to what students have recently learned
- Provide reports in a timely manner

Then you can…

- Address some of the disparities in background knowledge
- Provide more frequent and timely feedback to students and instructors
- Create space for course corrections
- Support teachers in planning instruction & scaffolding material
- Measure the acquisition of knowledge more effectively

Which will…

- Improve student experience and outcomes
- Make assessments more equitable
- Improve coherence between instruction, curriculum and assessments

Improve student learning

The research supports our hypothesis: Through-year assessments that connect more closely to what students are taught have the potential to improve student learning

By providing timely feedback to students

By creating space for course corrections

By creating greater coherence between instruction, curriculum and assessments

By increasing the acquisition and retention of knowledge

By decreasing the role of background knowledge in student performance

*This includes **curriculum-aligned** and **curriculum-relevant** through-year assessments*

Student learning improves when students are given timely and relevant feedback

What the research says

Students learn best when they receive timely and relevant feedback. Research shows that timeliness of feedback is critical to its effectiveness in order for it to resonate and impact a students' next task.

Claim

If the reports provided by curriculum-aligned and curriculum-relevant through-year assessments provide timely formalized feedback for students and teachers, then student learning will improve.

Sources: Karaman (2021), Wiggins (2012), Dawson et al (2019)

Student learning improves when teachers make course corrections based on their needs

What the research says

Course corrections are when an instructor uses the information gained from an assessment to inform and change their instruction. This could mean providing acceleration or reshuffling a scope and sequence. These kinds of behaviors can support student learning by increasing the personalization of the student experience. Instruction that adapts to individual students to help them access grade level instruction can improve student learning.

Claim

Curriculum-aligned and curriculum-relevant through-year assessments provide more frequent and timely reports of student performance. If teachers use these reports to scaffold instruction for students to access grade level instruction, student learning will improve.

Sources: Zhou et al (2022), Krupat & Dienstag (2009), Davis & Karunathilake (2004)

Student learning improves with increased coherence between curriculum, instruction and assessments

What the research says

When assessments are integrated into coherent systems that include high-quality curriculum and rigorous instruction, and are moving towards a united goal of improving student outcomes, student learning can improve as a result. Coherence with curriculum often more sharply connects with student learning than standards because it is tied to specific content.

Claim

If curriculum-aligned and curriculum-relevant through-year assessments increase coherence between instruction, curriculum and assessments, then student learning will improve.

Sources: Shepard et al (2017), National Academies Press (2001)

Student learning–and in particular, the acquisition and retention of knowledge–improves with increased opportunities to retrieve information

What the research says

The act of retrieving information helps improve student learning by increasing the acquisition and the retention of knowledge. An assessment requires students to retrieve information.

Claim

If curriculum-aligned and curriculum-relevant through-year assessments increase the number of times student must retrieve information they've learned, then the acquisition and retention of knowledge will increase, and student learning will improve.

Sources: Roediger III (2014)

Through-year assessments that connect to what students are taught can potentially reduce the impact of disparities in background knowledge on student performance

What the research says

Background knowledge plays a significant role in reading comprehension. Students do not have equitable access to opportunities to develop background knowledge and often are penalized for this when confronted with cold reads that are included in traditional end-of-year summative assessments.

Claim

If curriculum-aligned and curriculum-relevant through-year assessments include material students have seen before, which can decrease the impact of inequitable access to building background knowledge on a student's performance, then student learning will improve (and become more equitable).

Cold reads: texts that students have not read before and are topically unrelated to material taught in class

Warm reads: texts students have not read in class but are topically related to the information and knowledge they encountered in class

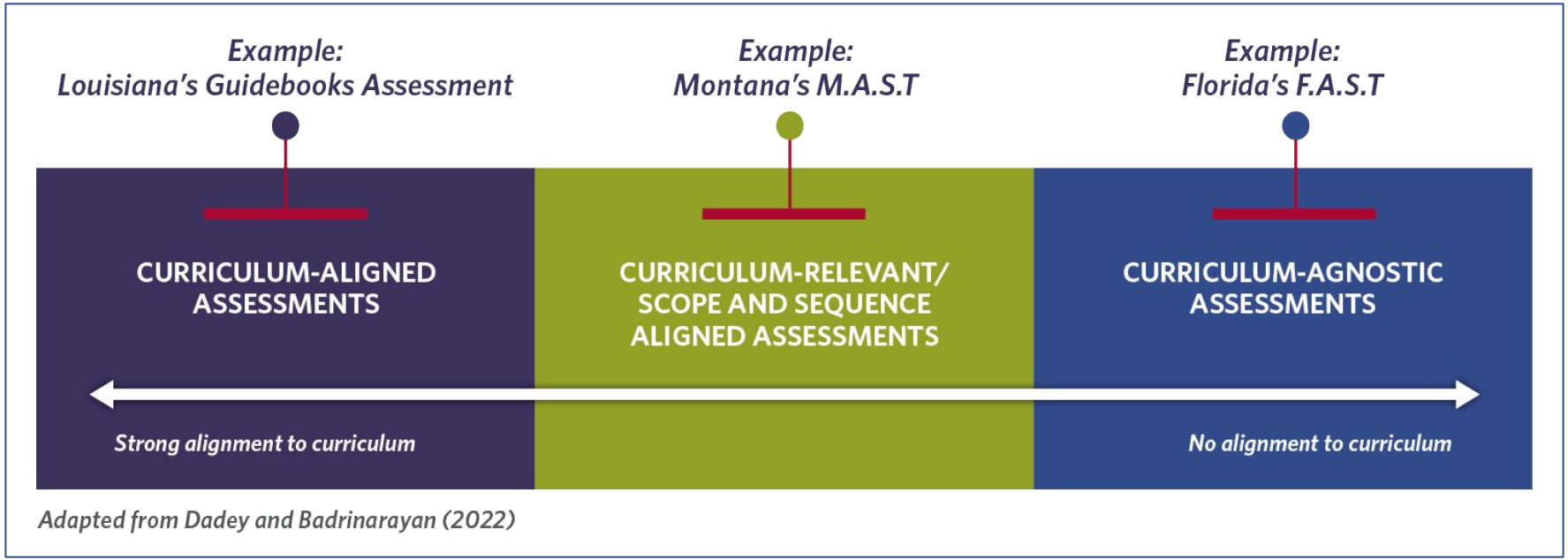
Hot reads: texts students have read in class

Sources: Smith et al (2021), Cromley & Azevedo (2007)

Tensions and Trade-Offs

Different through-year assessment approaches leverage the research differently and connect to instruction in varying ways

Through-year Approaches



Adapted from Dadey and Badrinarayan (2022)

Within these three approaches, there are still some tensions and trade-offs that states and assessment developers are working through

Alignment to Curriculum and/or Scope and Sequence:

What is the balance between local control and state assessments?

Timing:

When are students expected to demonstrate mastery of the content they've learned?

Scoring:

Should states measure the retention of knowledge or the acquisition of knowledge?

As through-year assessments become more tightly tied to a curriculum or scope & sequence, both the potential benefits and the logistical challenges may increase

Assessments **more tightly tied** to a curriculum or scope & sequence

The likelihood of **students being tested on material they have learned** in a timely manner

The **logistical challenges** of ensuring students have been taught material before they're tested on it

Curriculum-relevant and curriculum-aligned approaches have potential benefits depending on the level of flexibility they provide and how closely they are aligned with curriculum

Curriculum-relevant approaches

- The ability to transfer the assessments to different curricula which may make scaling statewide potentially more feasible

- Students will be tested on what they have been taught
- Students will be tested in a timely manner
- Instructors will be able to receive actionable data for instructional use

Curriculum-aligned approaches

- The ability to test students on specific material/texts they've been taught
- The ability to minimize the impact of previous background knowledge on test performance

On the other hand, both approaches also surface tensions related to implementation and local control depending on how closely they are aligned with curriculum

Curriculum-relevant approaches

- Districts need flexibility in the timing of when assessments are administered in order to meet their curriculum needs
- If this flexibility is not reached, then students may be tested on material they have not learned yet

- HQIM needs to be followed with fidelity, otherwise students will be tested on materials they have not learned yet

Curriculum-aligned approaches

- Data is needed on what curriculum instructors are using and the fidelity with which instructors are following the curriculum
- Assessments are not easily transferable to other districts not using the same curriculum

States vary on whether they test students on all standards during each administration or a subset of standards

	Each administration tests students on *the depth and breadth* of the state standards	Each administration tests students on *a subset* of the state standards
Potential Benefits	The model can be used in all schools within a state, regardless of local curriculum or instructional pacing	Students will be tested on material they have recently been taught
Tensions	Students may be tested on standards that haven't been taught yet	State testing is meant to measure mastery by the end of the year rather than proficiency at the time of testing
States (*of the eight identified*)	Louisiana, Florida, Nebraska, North Carolina (ELA), Texas	Montana, Delaware, Indiana, North Carolina (Math)

Dadey, N., & Gong, B. (2023) explore these designs in more depth.

States developing through-year models have different perspectives on whether data should be used to inform instruction, predict and/or contribute to summative scores

Purpose of the Data from Assessments

To inform instruction	To predict how students will perform at the end of the year	To contribute to summative scores for accountability

Note that these purposes are not mutually exclusive

States' varying perspectives also drive their design decisions

State Goal	Design Decision	
+ Measure acquisition of knowledge and use scores to inform instruction. + Students are given multiple opportunities to demonstrate their mastery of knowledge	+ Use a blueprint that is aligned to what students have been taught immediately before administration + Use aggregate scores from multiple administrations to calculate summative	 *For example:* Montana
+ Measure the retention of knowledge and use scores to predict how students will perform on the final summative test + Students are only given one opportunity to demonstrate their mastery of knowledge	+ Use a blueprint that incorporates the full scope of the standards during each administration + Use end of year administration for calculating summative	 *For example:* Nebraska

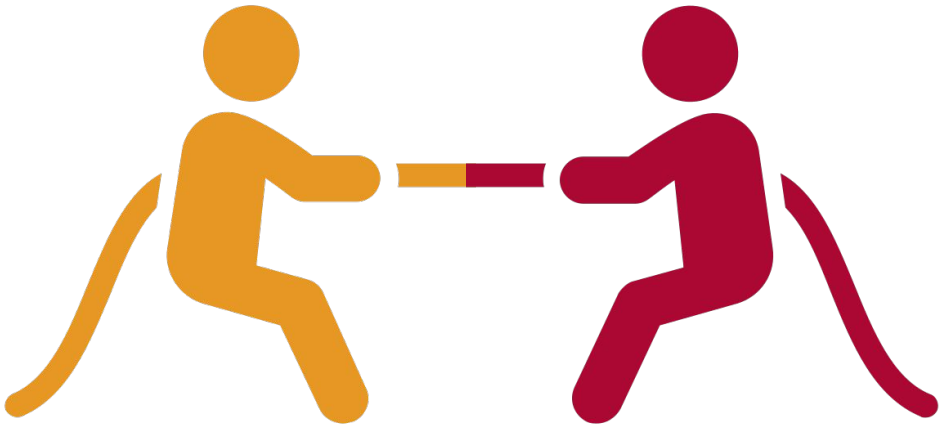
Dadey, N., & Gong, B. (2023) explore these design decisions in more depth.

How a state calculates their summative score also underlines fundamental beliefs about what we expect of students

States, members of measurement community, instructors, caregivers, etc

Some believe…

- Students should demonstrate mastery of standards after they've been taught the content
- Acquisition of knowledge and growth throughout the year is what we care about
- Accountability should drive improvement throughout the year



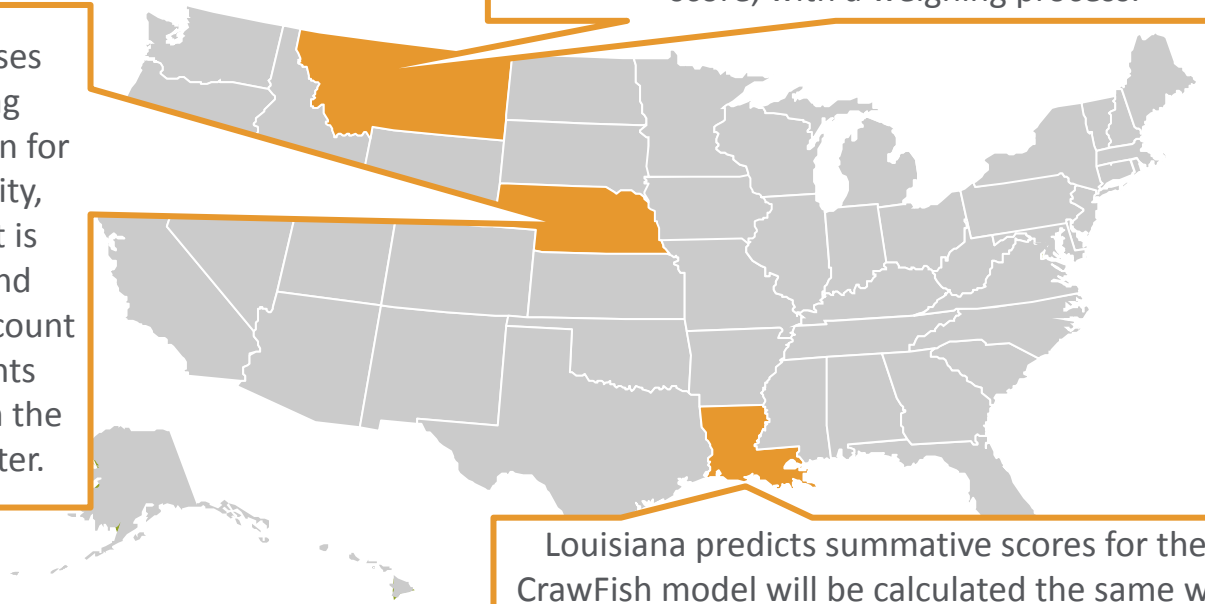
Others believe…

- Students should demonstrate mastery of standards at the end of the year
- Retention of knowledge is what we care about
- Accountability plus interims might distort implementation

Overall, many states are still figuring out how to calculate their summative score and there is a lot of variation among them

Montana eventually plans to aggregate mini-scale scores from each testlet into an overall summative score, with a weighing process.

Nebraska uses their spring administration for accountability, but the test is adaptive and takes into account how students performed in the fall and winter.



Across the states we reviewed, there is no agreement on how to calculate scores and varying opinions on the implications of those different choices.

Louisiana predicts summative scores for the CrawFish model will be calculated the same way as their curriculum-aligned models–pooling data to estimate scale scores.

While some of these tensions and considerations are unique to through-year models, many would be true for any approach that deviates from a traditional summative test

The summative testing infrastructure in the United States' education system is built around end-of-year assessments.

Changing this system will require more effort and reexamining of fundamental aspects of what is measured, for what purpose and how the data is used.

There are key enabling conditions and lessons learned from states testing these approaches that are worth learning from.

APPENDIX

Acknowledgements

We want to thank the following individuals for their time and expertise and sharing their learnings with us:

- **Elizabeth Gilbert**, **Jim Mirabelli**, **Laura Slover** and **Melissa Peplinski,** CenterPoint Education
- **The Center for Assessment team**
- **Dusty Shockley** and **Michael Feldman**, Delaware Department of Education
- **Sheila Briggs**, **Amanda Kim, Fatima Omar** and **Jennifer Vranek,** Education First
- **Vince Verges**, Florida Department of Education
- **Lynn Schemel**, Indiana Department of Education
- **David Steiner,** Johns Hopkins University
- **Thomas Lambert**, Louisiana Department of Education
- **Krystal Smith** and **Sharyl Allen**, Montana's Office of Public Instruction
- **Trudy Clark**, Nebraska Department of Education
- **Ashley Eden, Kristopher John** and **Chris Gentile,** New Meridian
- **Tammy Howard** and **K. Maxey-Moore**, North Carolina Department of Public Instruction
- **Leslie Mugan,** NWEA
- **Elizabeth Davis,** Odell Education
- **Iris Tian**, **Jamie Kwan** and **Jose Rios**, Texas Education Agency

This work is supported by the Bill & Melinda Gates Foundation, the Walton Family Foundation and the Chan Zuckerberg Initiative. Learn more about the next generation of curriculum-connected through-year assessments here.

Sources

Cromley, J. & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. Journal of Education Psychology, 99(2), 311-325. doi:10.1037/0022-0663.99.2.311\

Dadey, N., Evans, C. M., & Lorié, W. (2023). Through-year assessment: Ten key considerations. The National Center for the Improvement of Educational Assessment. https://www.nciea.org/wp-content/uploads/2023/03/Ten-Key-Considerations-Through-Year-Assessment-Report-March2023-F.pdf

Dadey, N., and Badrinarayan, A. (2022, April 21). In Search of the "Just Right" Connection Between Curriculum and Assessment: Considering Options Between Curriculum-Specific and Curriculum-Agnostic State Assessment. Center for Assessment. https://www.nciea.org/blog/in-search-of-the-just-right-connection-between-curriculum-and-assessment/

Dadey, N., & Gong, B. (2023). An introduction to considerations for through-year assessment programs: purposes, design, development, evaluation [Research report]. Retrieved from Smarter Balanced website: https://portal.smarterbalanced.org/library/en/2023-sb-consideration-of-technical-issues.pdf

Davis, M. H., & Karunathilake, I. (2004). The adaptive curriculum. *Medical teacher,* 26(6), 501–503. https://doi.org/10.1080/01421590412331285388

Sources

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2018). What makes for effective feedback: Staff and Student Perspectives. Assessment & Evaluation in Higher Education, 44(1), 25–36. https://doi.org/10.1080/02602938.2018.1467877

Education First. (2022). What are through-year assessments?: Exploring multiple approaches to through-year design. Retrieved from https://www.education-first.com/wp-content/uploads/2023/01/What-are-Through-year-Assessments-1.pdf

Karaman, P. (2021). The effect of formative assessment practices on student learning: A meta-analysis study. International Journal of Assessment Tools in Education, 8(4), 801–817. https://doi.org/10.21449/ijate.870300

Krupat, E., & Dienstag, J. L. (2009). Commentary: Assessment is an educational tool. *Academic medicine : journal of the Association of American Medical Colleges*, 84(5), 548–550. https://doi.org/10.1097/ACM.0b013e31819f7fb9

Marion, S. (2021, September 29). Trying to Serve Multiple Uses with Through Year Assessments. Center for Assessment.

Roediger, H. L. (2014). How tests make us smarter. The New York Times.

Sources

The National Academies Press. (2001). Implications of the New Foundations for Assessment Design. In *Knowing What Students Know: The Science and Design of Educational Assessment* (pp. 177–220). chapter.

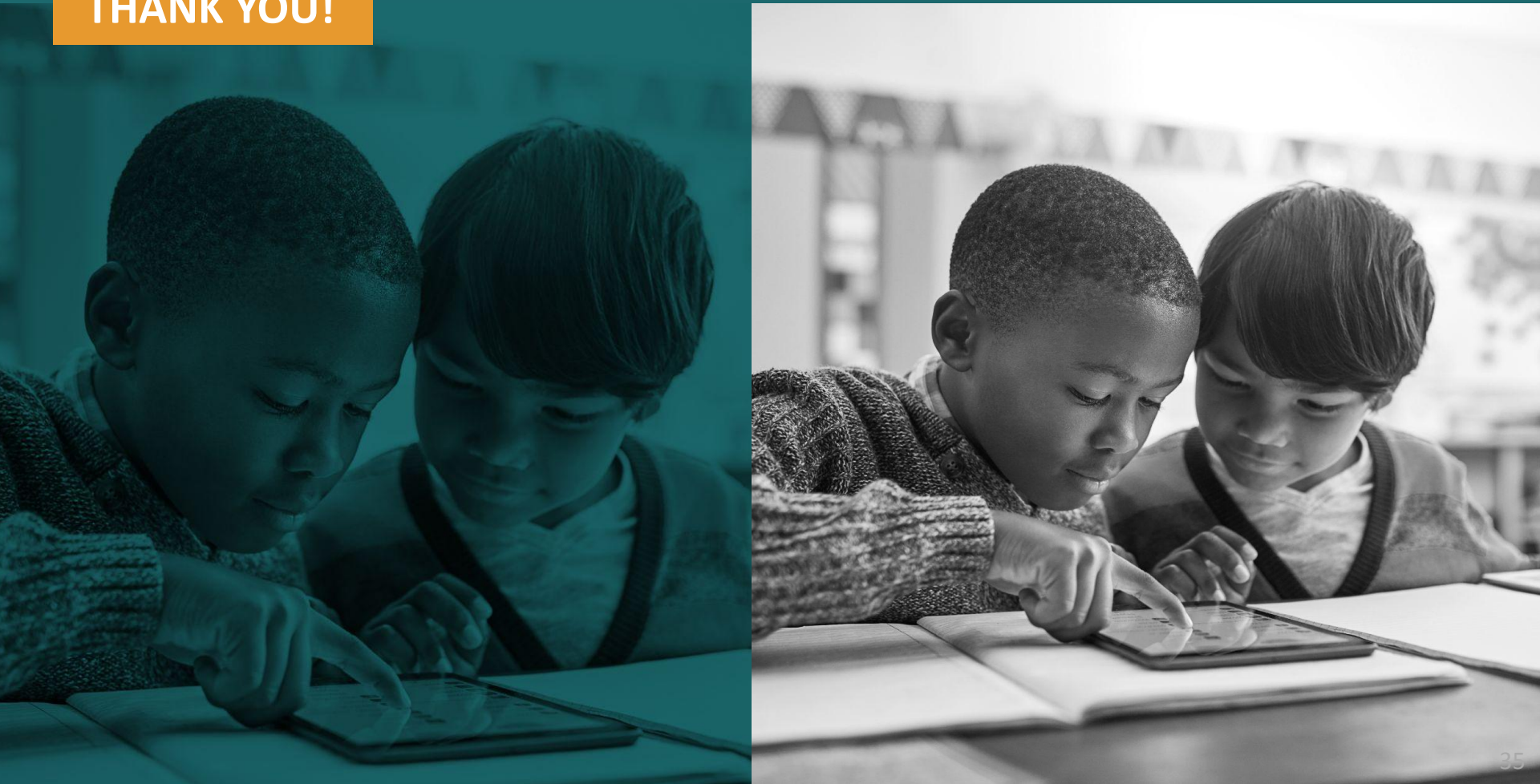
Shepard, L. A., Penuel, W. R., & Davidson, K. L. (2017). Design principles for new systems of assessment. *Phi Delta Kappan, 98*(6), 47–52. https://doi.org/10.1177/0031721717696478

Smith, R., Snow, P., Serry, T., & Hammond, L. (2021) The Role of Background Knowledge in Reading Comprehension: A Critical Review, Reading Psychology, 42:3, 214-240, DOI: 10.1080/02702711.2021.1888348

Wiggins, G. (2012, September). Seven keys to effective feedback. Educational Leadership, 70(1), 10-16. Retrieved August 22, 2023, from http://www.ascd.org/publications/educational-leadership/sept12/vol70/num...

Zhou, Y., Ye, X., & Liu, Y. (2022). The influence of personalized learning intervention system on student learning a study of junior middle school. *Interact. Technol. Smart Educ., 19*, 441-459.

THANK YOU!



Want to learn more?

Executive Summary

Why Through-Year?

Lessons Learned

Full Report

Exploring the Models

Return to the landing page