

# POSITIONING STATE ASSESSMENT SYSTEMS IN SERVICE TO TEACHING AND LEARNING:

## THE ROLE OF HIGH-QUALITY CURRICULUM IN STATE ASSESSMENT DESIGN

Authors:

Aneesha Badrinarayan and Dr. David Steiner

# TABLE OF CONTENTS

Forward.....	<b>3</b>
Introduction.....	<b>5</b>
What are curriculum-anchored assessment systems? .....	<b>6</b>
What are the design considerations and affordances of different approaches to curriculum-anchored assessment systems?.....	<b>11</b>
How can states design curriculum-anchored assessment systems that balance federal requirements for state assessments and local curriculum decisions? .....	<b>25</b>
What are the enabling conditions in states?.....	<b>34</b>
Conclusion .....	<b>36</b>
Authors & Acknowledgements .....	<b>39</b>
Sources .....	<b>40</b>

Front and back cover images by Allison Shelley/The Verbatim Agency for EDUimages.





**By design, current summative assessments used by states are disconnected from curriculum and instruction. This choice is driven by a goal of impartiality: to create a level playing field by ensuring none of the students are familiar with the specific texts and tasks on the test. The disconnect from curriculum, however, ends up reinforcing inequity and reducing instructional utility.**

**W**hen students sit for a test, they bring with them various experiences and prior knowledge. Since students with privileged backgrounds, such as students from the dominant culture or wealthier families and native English learners, are more likely to be exposed to the background knowledge necessary to interpret test items, these tests generate results that reflect societal inequities.<sup>1</sup> While informing instruction has historically not been the purpose of summative assessments, this exacerbates the challenge of utilizing them for instruction since it becomes hard to know whether variations in student performance reflect differences in prior knowledge or learning. Assessments' disconnect from curriculum make the results less meaningful for educators in other ways as well, for example by making it difficult to apply findings in the data to the daily instruction guided by the curriculum.

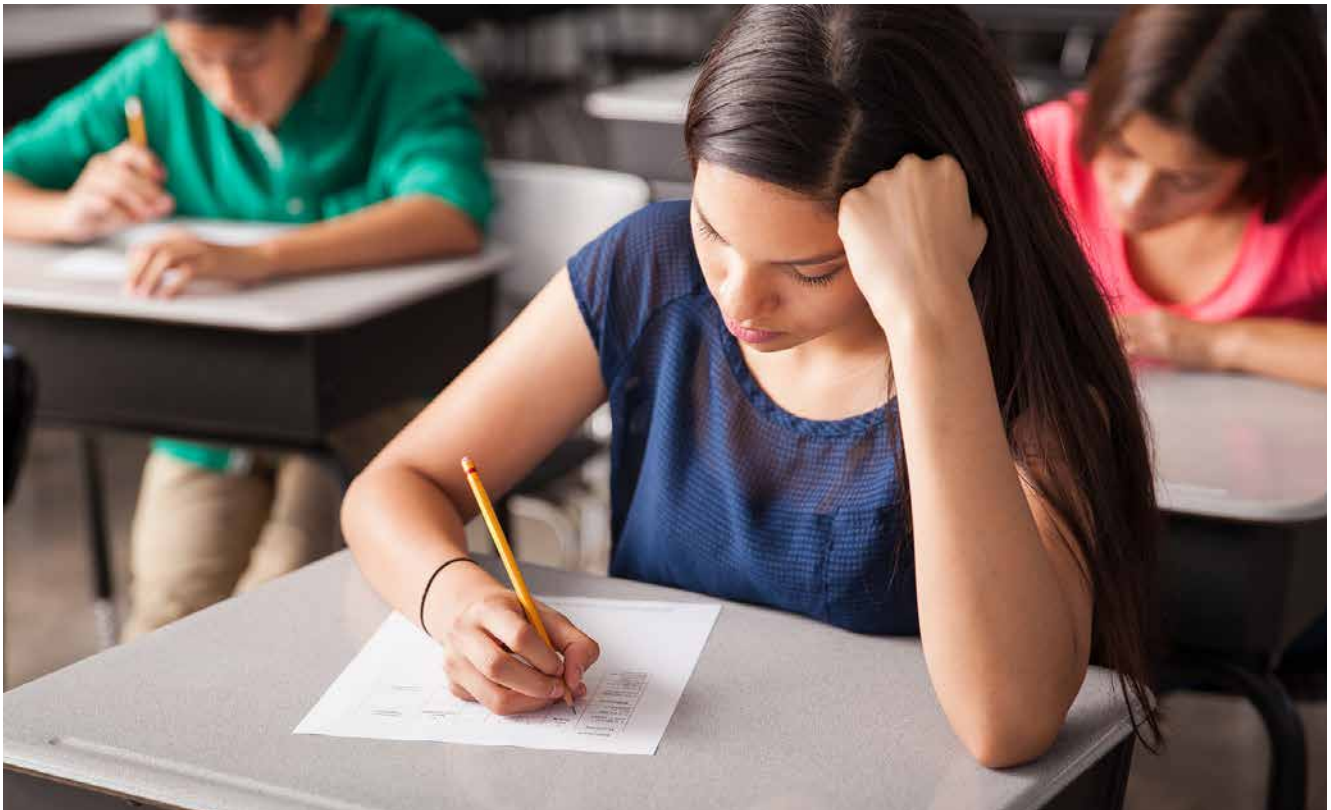
At Education First, we believe anchoring assessments in curriculum creates potential for more equitable, relevant and coherent state assessments that serve teaching and learning in addition to system evaluation. Through the **through-year curriculum-connected assessment grant program**—funded by the Bill & Melinda Gates Foundation, the Walton Family Foundation and the Chan Zuckerberg Initiative<sup>2</sup>—we are partnering with assessment developers and state education agencies

<sup>1</sup> For further discussion on the impact of prior knowledge on assessments, see Willingham, D. T. (2017) and Wexler, N. (2019).

<sup>2</sup> This paper is part of a series of products we are publishing to aid the field in understanding curriculum-connected, through-year assessments, share what we are learning during the prototyping and document how states and developers are grappling with challenges. Our first paper surveys through-year assessments being developed and piloted across the country and the different design choices states and assessment developers are making. We also have created a toolkit to support state leaders considering or in the process of transitioning to a through-year assessment system.

to explore and test this potential by developing and prototyping a new generation of state assessments. The approach brings together two innovations in summative assessments: through-year assessments and curriculum-connection. Through-year assessments measure student performance multiple times over the school year through periodic assessments instead of during a single sitting close to the end of the year. This process allows educators, families and students to receive data on performance when it can inform teaching and learning. To ensure this data is accurate and actionable, it is essential through-year assessments are also anchored in high-quality curriculum. Connecting the assessments to curriculum may reduce inequity generated by variations in background knowledge by grounding the assessment in the common set of knowledge and content students are exposed to in the curriculum. It may also make the data more useful for informing instruction by allowing educators to interpret and act on the data within the context of the specific content students are engaging in. Lastly, curriculum-connection helps address one of the fundamental challenges of through-year assessments—ensuring students are exposed to the content they are tested on *before* sitting for periodic assessments.

Connecting state assessments to curriculum, however, creates its own distinct challenges. The wide variety of curricula utilized within a state can make it difficult to envision how a statewide test can be anchored in curriculum. We hope this paper will seed thinking about how design choices can alleviate this challenge in order to develop state assessments that realize the benefits of curriculum-connection while also meeting federal requirements.



# INTRODUCTION



## Introduction

As states face growing calls to develop assessments that more directly support high-quality teaching and learning, several states and education leaders are considering ways to connect state assessments to what students experience and learn in the classroom. When done well and implemented with care, states can use curriculum-anchored assessments to:

1. Learn more effectively,
2. Yield more valid and meaningful test scores, and
3. Have a positive influence on the trajectory of classroom instruction.

This paper explores a set of five key questions related to curriculum-anchored assessments, including:

- + **What** are curriculum-anchored assessments?
- + **Why** consider curriculum-anchored assessments?
- + What are the **design considerations and affordances** of different approaches to curriculum-based assessment systems?
- + **How** can states design curriculum-anchored assessment systems that balance federal requirements for state assessments and local curriculum decisions?
- + What are the **enabling conditions** for curriculum-anchored assessment systems in states?

Because curriculum-anchored assessment systems are still an emerging model in current state assessment systems, this paper considers examples, lessons learned, opportunities and challenges from an expansive set of curriculum-anchored systems. These systems include:

- + Robust systems used internationally,
- + High-quality systems in place in the United States prior to the No Child Left Behind era of testing,
- + Those currently used in states and districts as part of local and statewide efforts at innovative, instructionally relevant assessment systems, and
- + Large-scale integrated curriculum and assessment systems, like Advanced Placement and International Baccalaureate.

# DEFINING CURRICULUM-ANCHORED



## What are curriculum-anchored assessment systems?

Curriculum<sup>3</sup>-anchored assessment models intentionally connect assessments used to make judgements about students' growth and achievement to students' learning experiences in their classrooms.

There are many different ways to conceptualize how assessments can connect with curriculum<sup>4</sup> in terms of the content, timing, and supports for assessment use (Figure 1). For example, in systems that are designed to tightly couple curriculum and assessment, the texts, scenarios, contexts and questions students respond to on an assessment directly reference the specific content of a given set of instructional materials. In this case, student performance on the external assessment is predicated on students' engaging with those *specific* instructional materials in their classrooms. Other curriculum-anchored systems are designed for a looser coupling between curriculum and assessment, but these systems still work to create a connection between the two (e.g., using a common topic or theme to link together instructional materials and assessment content). This design allows the assessment to connect to a wider range of instructional decisions and materials.

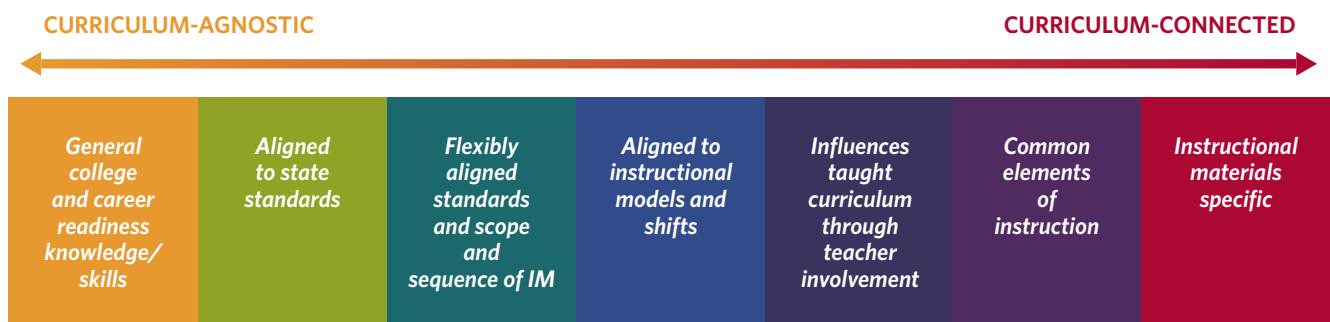


Figure 1. Spectrum of ways to “anchor” assessments to curricula

<sup>3</sup> In this paper, curriculum is used to refer to the full range of teaching and learning experiences students have throughout the course of learning within a given discipline and/or course. It is inclusive of all instructional materials as well as pedagogical approaches used by teachers. When a more specific component of curriculum (e.g., instructional materials) is the focus of a discussion in this paper, we seek to explicitly call that out.

<sup>4</sup> For further discussion, see Dadey, N. & Badrinarayan, A. (2022), whose work also influenced the framing and presentation of Figure 1 through 4

## DEFINING CURRICULUM-ANCHORED

In all cases, curriculum-anchored assessment systems explicitly acknowledge that there is a **reciprocal relationship between what and how students learn and how they perform on an external assessment**. Curriculum-anchored assessments seek to create instructional coherence by providing more accurate measures of what students know and can do while simultaneously reinforcing and incentivizing the systemic use of high-quality instructional materials and practices.

### Why are system leaders considering curriculum-anchored state assessment systems?

There is no such thing as a truly “curriculum-agnostic” state assessment. Although current state assessments attempt to side-step curriculum by focusing on end-of-instruction goals (i.e., state standards), the reality is that all assessments (1) make certain assumptions (intentionally or unintentionally) about what students will learn or experience in the classroom, (2) are sensitive to curricular opportunities to learn, which influence both the design of assessment questions and the interpretation of assessment data that follows, and (3) signal and to some degree incentivize districts’ and teachers’ curricular choices in an effort to maximize scores on the state assessment.

Given this relationship, there are at least three major reasons why many states are considering curriculum-anchored systems:

- + To improve the validity of test scores,
- + To position assessments as a force for better instruction,
- + To surface truly actionable information.

**Improving the validity of test scores:** Curriculum-agnostic assessment systems do not reflect what research and practice tell us about how to best support and measure learning. In an effort to be curriculum-agnostic, most state standards and their related assessments have focused on decontextualized, end-of-instruction goals. This often results in assessments that focus on skills (especially on ELA, social studies, math tests) or decontextualized knowledge (particularly on science and history tests), rather than on the integrated development and demonstration of content-infused conceptual understanding and disciplinary practices. This is problematic for two reasons:

1. Emphasis on skills or content in relative isolation reinforces teaching practices that focus on practicing those skills or memorizing that content in relative isolation. This emphasis directly opposes what we know from research in the learning sciences: students develop expertise and mastery in many of the major disciplines—including English, social studies, and science—by developing conceptual understanding and deep content knowledge together with disciplinary practices and skills.<sup>5</sup> Furthermore, current tests press students to memorize isolated “factoids”

---

<sup>5</sup> It is worth recognizing that the vast majority of students taking state assessments have not yet mastered grade-level standards. If assessments are to be useful from an instructional perspective, it is particularly important that they are sensitive to what learning looks like along the way to proficiency and mastery. Current state assessments often confuse many interest holders by presenting performance and achievement level descriptors as a continuum of performance (e.g., basic/advanced) when they are not in fact describing students along a progression of learning.

## DEFINING CURRICULUM-ANCHORED

pulled out of nowhere. Once again, neuroscience and teaching experience tells us that this is the least effective way to facilitate the transfer of knowledge into long-term memory—a key operation to retain knowledge. In short, by focusing on skills or content in ways that are decoupled from high-quality teaching and learning approaches, state assessments often incentivize using ineffective teaching practices and instructional materials in a misguided effort to maximize scores on the test.

2. Scores on curriculum-agnostic assessments are often difficult to interpret (and may in fact misrepresent student understanding and ability), because what and how students learn directly influence how they perform on an assessment. This quickly becomes an issue of equity. A growing body of research suggests that students' relevant content knowledge influences how well they can make sense of texts—including those on the assessment. Curriculum-agnostic assessments give text passages that students have not studied previously, and so more privileged students will bring all the advantages of greater background knowledge to the questions about these passages. On such an assessment, it is difficult to determine whether poor student performance and score-gaps between students are due to a lack of relevant content knowledge or to specific vocabulary or comprehension difficulties with the assessment text. Either way, these results would be interpreted as a failure to demonstrate English learning targets. This result is problematic both because (1) the scores are not valid across all students (poor performing students may be able to demonstrate considerable proficiency on the same target as their higher performing colleagues if given texts about content they are equally familiar with) and (2) the reports do not point to the root cause for the difference in scores, which limits how well teachers and systems can support growth and learning acceleration. Thus curriculum-agnostic English language arts assessments tell us a great deal about students' opportunities to develop background knowledge related to topics that appear on state tests, but these tests do not act as effective measures of students' learning in their classrooms. Similar arguments can be made in social studies as well as science.

Put plainly, assessments are often sensitive to many elements of students' lives and learning experiences—including their prior knowledge, lived experiences and access to resources—whether we design for them or not. Without explicitly designing assessments to attend to what and how students learn, it can be difficult to understand how to interpret and use test scores—and indeed, whether test scores are actually representing what students truly know and can do fairly.

**Positioning assessments as a force for good instruction:** Given the interdependence between what is assessed and what students experience in the classroom, many states are beginning to ask how they can use assessments to directly drive positive shifts in teaching and learning. One way to do so is to imbue assessments—the tasks, the implementation and the reports—with discerning features of high-quality curriculum. This change makes assessments not only tests simply worth teaching “to” but also a tool to educate and guide leaders, teachers, and student development. This shift is not a new idea—indeed, many scholars have discussed the signaling power of assessments<sup>6</sup>, and in many countries around the

<sup>6</sup> For example, see Haertel, E. (2013).



## DEFINING CURRICULUM-ANCHORED

world, assessments intentionally carry with them an implied curriculum that is often set at a national or regional level. Even in the United States, there are already several large-scale assessment systems—such as interim assessments and the assessment systems embedded within Advanced Placement and International Baccalaureate programs—that intentionally build from and shape curriculum experiences. These systems attend to what, when, and how students learn to create assessments that better reflect and support teaching and learning.

One specific way assessments can be a force for good instruction is by **motivating evidence-based strategies to improve teaching and learning—namely, the use of high-quality instructional materials**. A compelling body of research suggests that teachers’ use of high-quality instructional materials (HQIM) is one of the most powerful determinants of student performance.<sup>7</sup> This effect is even more pronounced when the use of HQIM is done continuously and with integrity and when policies and resources (e.g., those for professional learning) at the school, district, and/or state level coherently support their uptake and implementation. The good news is that there is growing commitment from state leaders to actively support the availability, adoption and meaningful use of HQIM.<sup>8</sup>

Curriculum-anchored state assessment strategies could play an important role in state instructional materials and professional learning strategies. By explicitly connecting the state assessment to high-quality instructional materials, states can simultaneously encourage the **uptake** of high-quality instructional materials while also providing additional resources and support for the **implementation** of those materials with integrity. For example, resources that states devote to developing state assessment items, scoring open-ended responses, and other assessment professional learning activities can function as curriculum-based professional learning in states that connect their state assessments to high-quality instructional materials in some explicit way.

**Surfacing truly actionable information by directly linking student performance to the work students and teachers are doing in the classroom:** The distance between what is tested and what is taught, in terms of content and approach as well as timing, is one of the most significant barriers to current state assessments being considered instructionally relevant. Curriculum-anchored assessment systems address these concerns by directly connecting assessment results to teaching and student learning. While this has obvious implications for teachers, there are important implications for leaders as well. As illustrated in Table 1, adapted from work led by Aneesha Badrinarayan, Nathan Dadey and colleagues, leaders at every level of the system make decisions based at least in part on large-scale assessment data that influence how teaching and learning happen in the classroom. By providing data that are intimately linked with classroom activities, leaders can also make better decisions that are more likely to help students and teachers thrive.

<sup>7</sup> For example, see Grissmer, D., et al. (2023).

<sup>8</sup> See the CCSSO Instructional Materials and Professional Development Network.

# DEFINING CURRICULUM-ANCHORED

*Table 1. Instructionally Relevant Decisions and Actions Throughout the System*

INTEREST HOLDER	EXAMPLES OF INSTRUCTIONALLY-RELEVANT DECISIONS AND ACTIONS
Students	<ul style="list-style-type: none"> <li>+ Developing and reflecting on learning goals (metacognition).</li> <li>+ Self-monitoring progress toward learning goals.</li> <li>+ Co-designing learning experiences that are relevant and meaningful.</li> <li>+ Actively engaging in disciplinary inquiry.</li> <li>+ Being motivated to engage with instructional materials.</li> </ul>
Teachers	<ul style="list-style-type: none"> <li>+ Teaching content with care and attention.</li> <li>+ Engaging in culturally and linguistically responsive teaching practices.</li> <li>+ Eliciting and leveraging students' current understanding and experiences as a foundation to anchor new learning.</li> <li>+ Providing opportunities for deep, sustained and compelling learning.</li> <li>+ Using teaching strategies that encourage individual and social sense-making and academic risk taking.</li> <li>+ Engaging students in learning that mirrors the authentic behaviors and conceptual development of the discipline.</li> <li>+ Selecting and adapting instructional activities/materials in response to individual, school, and community interests, needs, and priorities.</li> <li>+ Determining how and when students receive quality feedback related to learning goals.</li> </ul>
School and district leadership	<ul style="list-style-type: none"> <li>+ Scheduling decisions that enrich curricular opportunities for all learners (e.g., inclusive learning for emerging multilingual learners, providing sustained and coherent time for science and social studies in K5).</li> <li>+ Using observation protocols and educator coaching that reflect the major instructional shifts of the discipline, including the specific instructional materials used in the classroom as appropriate.</li> <li>+ Implementing equitable and non-restrictive grading policies.</li> <li>+ Adopting and implementing high-quality instructional materials.</li> <li>+ Investing in system-wide curriculum-based professional learning.</li> <li>+ Establishing responsive course options and pathways that create opportunities for learners.</li> </ul>
State leadership	<ul style="list-style-type: none"> <li>+ Developing and adopting coherent standards, instructional materials, assessment and professional learning policies for pre-service and in-service teachers.</li> <li>+ Incentivizing the use of high-quality instructional materials through funding options (e.g., reimbursing the purchase of HQIM or allowing use of allocated curriculum funding for professional learning if using high-quality OER materials).</li> <li>+ Tailoring state-offered professional learning to support key instructional shifts, in the context of specific instructional materials as appropriate.</li> </ul>

## CURRICULUM-ANCHORED DESIGNS



### What are the design considerations and affordances of different approaches to curriculum-anchored assessment systems?

Two of the most important considerations for curriculum-anchored assessment systems are content and timing: (1) what should the content of the assessment be, and how does this relate to curriculum? and (2) when should the assessments be administered such that they provide the highest value-add?

#### Content considerations for curriculum-anchored assessment systems

State decisions about appropriate assessment design should stem from disciplinary content considerations. The best test for science will look very different from English language arts', because how students learn and demonstrate understanding—and therefore, how curriculum is designed and implemented—is different in science than in English language arts. Below, we discuss key considerations and implications for assessment design in English language arts, science and mathematics.<sup>9</sup>

#### English language arts (ELA)

##### Rationale for curriculum-anchored assessment systems

Student performance in ELA is closely related to student knowledge of the content or subject discussed in a given text. HQIM in ELA is content-rich and intentionally focuses on building students' knowledge in the areas about which they are reading and writing. By connecting assessments to the topics and texts students have studied in the classroom, assessments are better able to measure what students know and can do in ELA. This is in contrast to skills-focused assessments that conflate student performance on ELA skills-based standards with a student's background knowledge on the subject of a given text. In turn, assessments anchored to curriculum can incentivize the use of HQIM that focuses on knowledge building and not just decontextualized skills.

<sup>9</sup> We have limited this discussion to ELA, math, and science as the disciplines with federally mandated assessments because they share common policy opportunities and barriers. Curriculum-anchored assessment systems are also deeply relevant in many disciplines within social studies—considerations for social studies are often similar to ELA and/or science, depending on the specific sub-discipline.

# CURRICULUM-ANCHORED DESIGNS

## Implications for assessment design

Assessments should require students to analyze texts they have studied and should demonstrate that they can build on that knowledge when confronted with a new text that offers new perspectives about the same knowledge domain. This likely requires that either (1) texts on an assessment are drawn from a common set of topics or specific texts that students must have experienced in curriculum or (2) students are able to choose the texts they want to reference in response to assessment prompts.

## What are some options for how states might develop curriculum-anchored assessment systems in ELA?

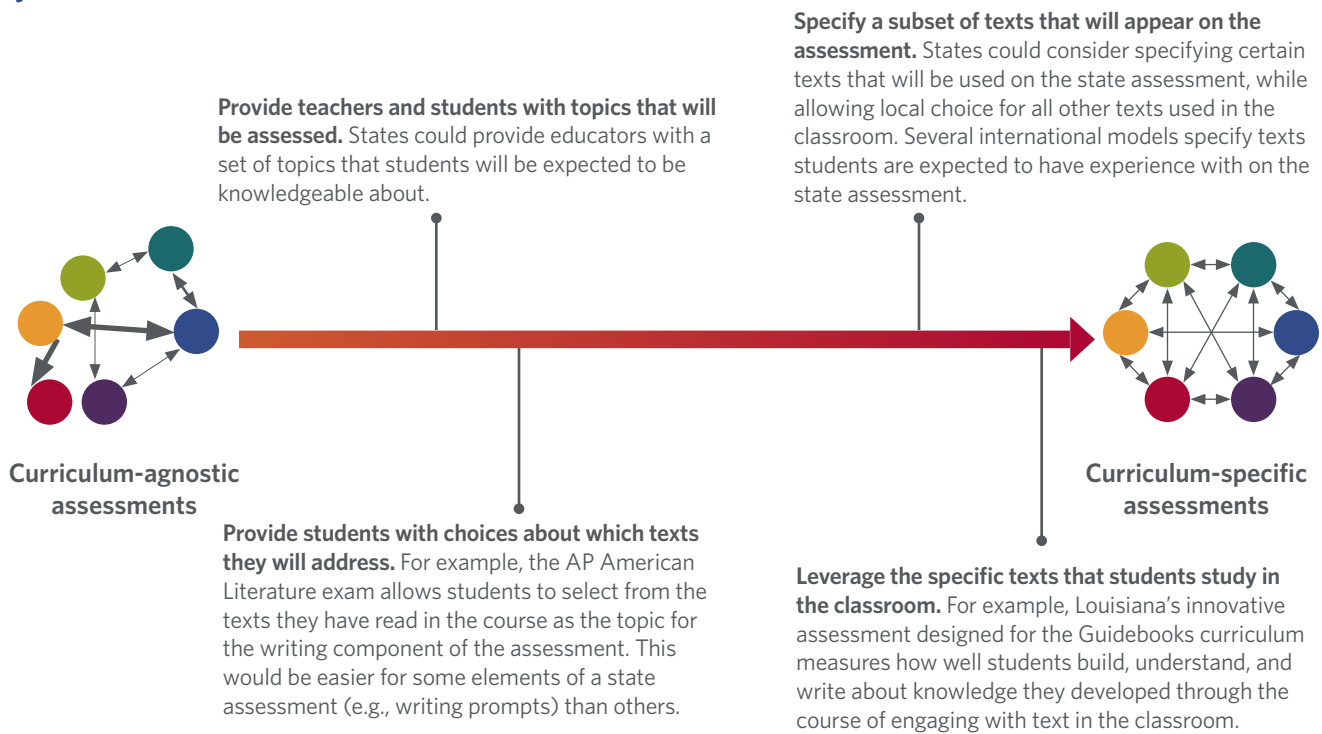


Figure 2: Spectrum of ways to “anchor” assessments to English language arts curricula

## Science

### Rationale for curriculum-anchored assessment systems

Over the last decade, nearly all states have adopted science standards that expect students to demonstrate achievement by making sense of phenomena and problems in the natural and designed world through the use of science ideas and practices. However, the majority of classrooms and schools have yet to realize this vision for teaching and learning, in part because science assessments do not yet embody the major implications for instructional shifts that are a core feature of teaching and learning efforts in science. Connecting assessments to high-quality curriculum ensures that assessments (1) attend to the distinct opportunities to learn created by

# CURRICULUM-ANCHORED DESIGNS

different phenomena, problems and sense-making approaches across high-quality curricula, (2) reflect the kinds of experiences students should be having in the classroom, both consistent with and complementary to HQIM, and (3) signal and incentivize the most important instructional shifts and performances in the discipline to teachers and leaders.

## What are the implications for assessment design?

Assessment designers will need to consider whether the purpose of curriculum-anchored assessments within a given context is (1) to incentivize the use of HQIM, (2) to support educators already using the materials or (3) to signal and support common elements of high-quality science teaching and learning that occur across multiple curriculum approaches. While the specifics of the models will vary based on the purpose, all assessments should be designed while explicitly attending to the phenomena, problems and approaches to learning—including sense-making routines developed, ways of using practices and science ideas together and important features of a multidimensional instructional model—students experience in high-quality science curriculum.

## What are some options for how states might develop curriculum-anchored assessment systems in science?

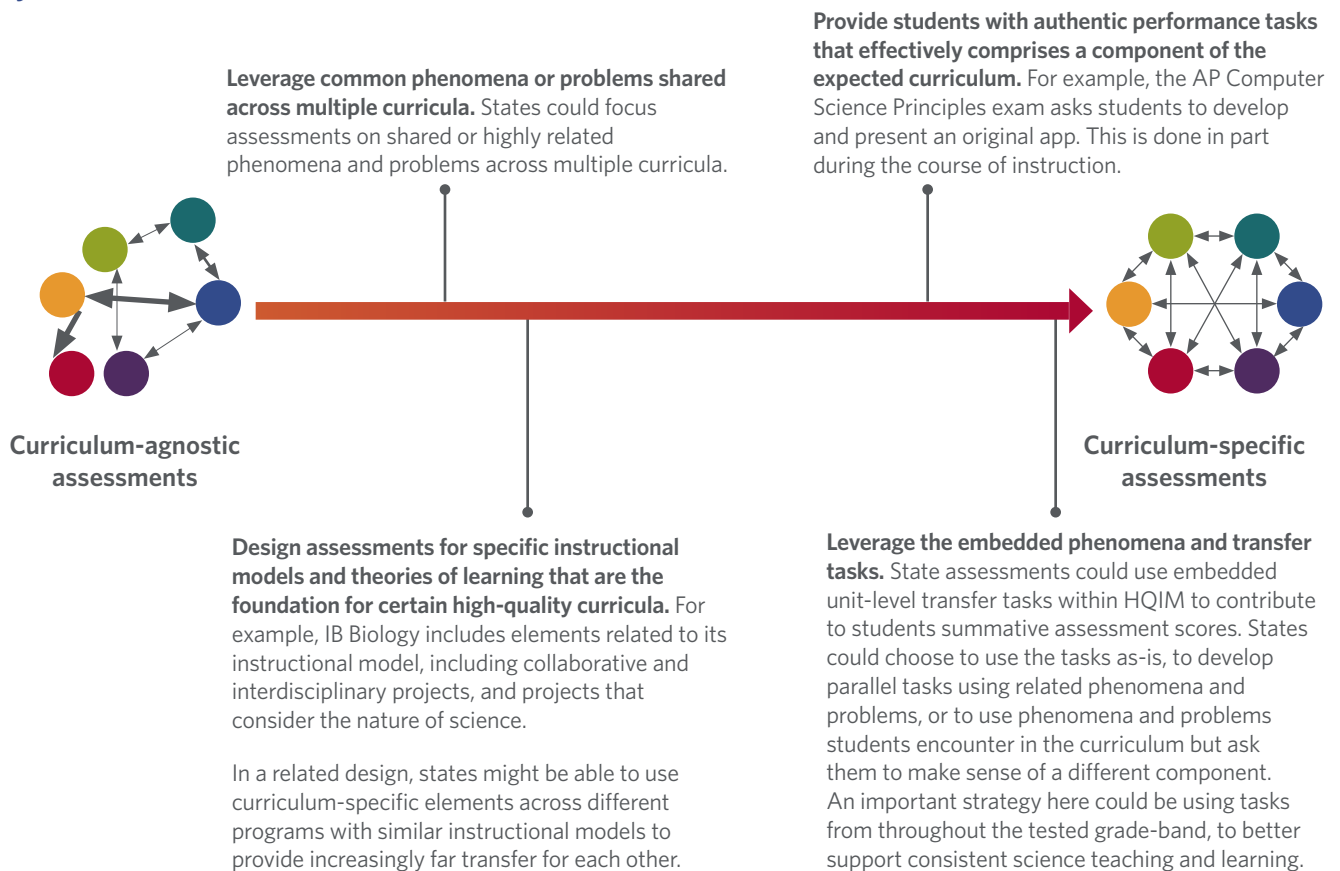


Figure 3: Spectrum of ways to “anchor” assessments to science curricula

# CURRICULUM-ANCHORED DESIGNS

## Mathematics

### Rationale for curriculum-anchored assessment systems

While many math curricula include the procedural elements of current math standards, high-quality teaching, learning, and instructional materials focus on developing students' conceptual understanding and ability to apply math concepts to real-world problems. Connecting assessments to high-quality curriculum in math incentivizes the use of curriculum that will emphasize these capacities while also providing more useful information about both student learning and effective teaching in mathematics, relative to specific curriculum.

### What are the implications for assessment design?

If the purpose of the assessment is to incentivize the use of high-quality curriculum, the assessment must be designed to send clear and discerning signals that the approaches taken in such curriculum are essential to success on the state math assessment. This may involve aligning to the specific scope and sequence of the materials, attending to specific attributes of how a set of materials intentionally builds students' understanding or focusing on problems/contexts embedded within HQIM.

### What are some options for how states might develop curriculum-anchored assessment systems in mathematics?

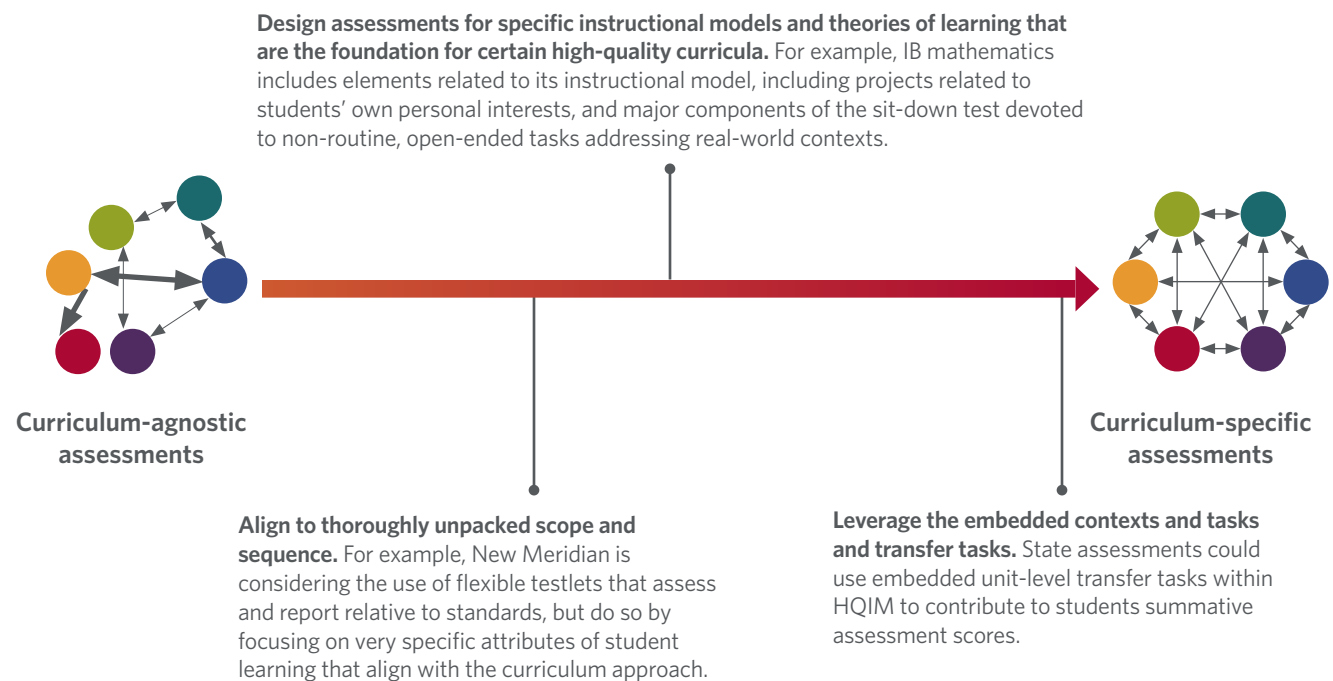


Figure 4: Spectrum of ways to “anchor” assessments to math curricula

## CURRICULUM-ANCHORED DESIGNS

In addition to disciplinary considerations, there are two important content-related considerations that are important to note across all disciplines:

**High-quality instructional materials.** All curriculum-anchored assessment systems must rest on the foundation of high-quality curriculum, including HQIM. These systems simply do not function if high-quality materials, teaching, and learning are not at their core. While the specific features of high-quality materials and curriculum will vary across disciplines, they are all grounded in research on how students learn and how that learning evolves over the course of K12 to ensure all students are supported in meeting state standards.

**Expectations of transfer and generalizability.** In every discipline, the ultimate goal of K12 teaching and learning is to prepare students to be able to build their disciplinary understanding and apply their developing understanding to new contexts as they may encounter them in college, career and citizenship while generalizing and updating their learning as appropriate.<sup>10</sup> Curriculum-anchored assessment systems assume that the best way not only to assess transfer but also to support its development is to (1) help students thoroughly develop disciplinary content, concepts and practices and (2) attend to both what students learned in the classroom and reasonably new contexts (e.g., texts, phenomena, problems, new connections across disciplinary ideas). For example, a curriculum-anchored ELA assessment may include:

- + Texts that students are very familiar with and may have recently studied,
- + Texts about topics in which students have developed content knowledge but that students haven't previously studied in their ELA curriculum,
- + Questions that ask students to synthesize their knowledge from across multiple texts and curricula units.

Triangulating student performance across these different contexts provides much richer information about how students can transfer understanding—information one can only surface by building the assessment on students' curricular experiences.

<sup>10</sup> It should be noted that how humans develop the ability to transfer ideas within and across different disciplinary contexts and real world situations is still a developing area of understanding. While many state assessment contexts assume that the mark of learning is that students can transfer or generalize what they learned in school to the questions on the test, research from learning and cognitive sciences suggests that this process is much more complex than assessment systems tend to assume. Many demonstrations of "transfer" actually often reflect that those students (1) have had more, and a more diverse range, of experiences that they can use to relate to a novel situation (in other words, the situation is "less novel" to some students than others), and (2) that this wider range of experience is coupled with deep disciplinary understanding that is cultivated through repeated opportunities to making sense of many related but different situations over time, such that students have enough information to apply and generalize their thinking. For more information, please see National Resource Council (2000), Healy, A. F., & Wohldmann, E. L. (2012), Wooldridge, C., & Weinstein, Y. (2016), and National Academies of Sciences, Engineering, and Medicine (2018).

# CURRICULUM-ANCHORED DESIGNS

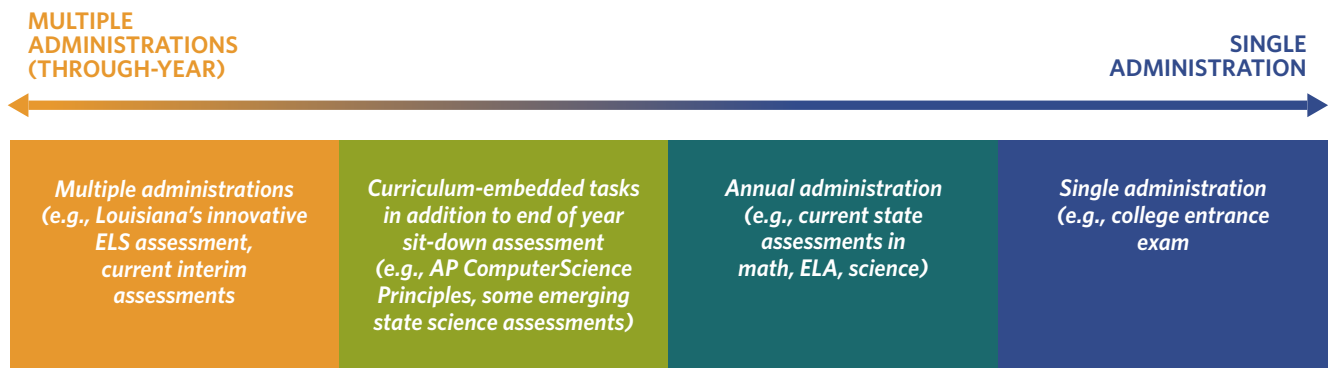


Figure 5. Options for the Timing and Cadence of Curriculum-Anchored Assessment Administration

## Timing and cadence considerations for curriculum-anchored assessment systems

All assessments, regardless of timing and cadence, can be curriculum-anchored. However, decisions about when and how often students are assessed can determine how curriculum-anchored assessments will be used.

Many states are considering through-year assessments<sup>11</sup> as one way to make state assessment systems more meaningful to students, teachers and families (Figure 5). While curriculum-anchored assessments can be administered once (e.g., AP exams, International A and O level exams) or multiple times a year (e.g., Louisiana's Guidebooks/Wit & Wisdom assessment, curriculum-anchored interim assessments), it is essential that through-year assessments be curriculum-anchored. Without carefully attending to what and how students are learning throughout instruction, through-year assessments can inadvertently magnify the same harm that current curriculum-agnostic assessments cause by misaligning assessments and instruction.

## What is gained and what is lost when curriculum-anchored assessment systems are administered once or more than once within a given school year?

Because others have extensively discussed considerations and limitations applicable to the possible use of through-year assessments,<sup>12</sup> this paper will focus on what is gained and what is lost when curriculum-anchored assessment systems are administered once or more than once within a given school year.

### Major curriculum considerations for periodic (e.g., through-year) assessments

- + Modular and progressive elements of the curriculum.** States and developers will need to consider which elements of the curriculum can be assessed during periodic assessments throughout the year and which elements of the curriculum build gradually over time and are thus best assessed at the end of the academic year.

<sup>11</sup> See Education First (2022).

<sup>12</sup> Our colleagues at the Center for Assessment have published and presented extensively on this topic. For particularly relevant work, please see Dadey, N., et al. (2023) and Dadey, N., & Gong, B. (2023).



## CURRICULUM-ANCHORED DESIGNS

- + **Range of curricular implementation.** Some curriculum materials are intended to be implemented along a set scope and sequence, while others are designed such that teachers, schools and districts can make decisions about the order and timing of unit delivery. Even in contexts where curriculum is supposed to follow a particular path, there may be variations. Designing through-year curriculum-anchored assessments requires that states be knowledgeable about the range of implementation efforts and approaches for a given curriculum and across curricula.
- + **Quality and most relevant instructional shifts.** Even more than a single end-of-year assessment, through-year assessments have the power to drive instruction. It will be imperative that states and developers (1) ensure through-year assessments are truly reflective of authentic learning and activities within the discipline and (2) design through-year assessments that signal and incentivize high-quality teaching and learning in the discipline, with a particular focus on ensuring that the assessments support educators in making the most needed shifts in instruction from current practice in the context of the use high-quality curricula.

### Affordances of periodic assessment administration for curriculum-anchored assessments

Assuming periodic assessments are designed well, some possible affordances of through-year approaches to curriculum-anchored assessments include:

- + **Improving teacher practice in real time.** Through-year assessments may incentivize teachers to teach the adopted instructional materials. They also give teachers real-time information about students' developing knowledge and practice that enables teachers to adapt their teaching practices within a given school year and thus benefits the students taking the assessments. In some content areas and with some curricular approaches, this feedback may also be useful to support student learning. One limitation is that, depending on the design of the curriculum (e.g., what content and practice is revisited vs. what is addressed at a given moment within instructional materials) there may not be time to revisit student needs and still prepare students for the next testing window.
- + **Deep probing of a given unit.** Through-year administrations can allow assessments to measure depth and degree of student understanding, provided that the timing of these assessments does not overly interfere with instruction. In some content areas, deep probes of a unit may also elevate evidence of student progress in areas that are more difficult to probe in an end-of-instruction assessment. For example, in Louisiana's innovative assessment model, assessments are able to go beyond reading comprehension and surface evidence of students' ability to synthesize across texts they worked with during the associated instructional unit. In science, this might mean more deeply assessing students' developing ability to transfer understanding from the specific phenomena or problems addressed in the unit.
- + **Motivating student practice.** Through-year assessment administration can give students a reason to practice cognitively demanding knowledge and abilities within the curriculum, such as synthesis and transfer, that require sustained engagement over time to cultivate.

## CURRICULUM-ANCHORED DESIGNS

- + **Encouraging the use of high-quality formative processes.** By introducing assessments with stakes more routinely throughout the year, it is possible that districts, schools, and teachers may use better formative assessment practices. If designed to center meaningful instructional shifts and be educational to teachers and students, through-year assessments may themselves serve some of these purposes—thus streamlining the assessments used in classrooms.
- + **Monitoring student progress alongside the specific development of practice and schema.** Through-year assessments can be sensitive to the specific ways and degrees students are developing content and practicing knowledge.

### Affordances of more limited assessment administration for curriculum-anchored assessments

While periodic assessment administration may offer more timely information for teachers, there are several potential benefits to providing curriculum-anchored systems that are administered less frequently (e.g., end of year, flexible performance tasks administered during the year). Some of the benefits of a more conservative approach to testing frequency and cadence include:

- + **Allowing flexible and responsive teaching that is not tied to the sequencing of assessment administration.** While through-year models may motivate teachers to strive to stay on track with a curriculum's scope and sequence, a single annual administration (or more limited administrations, like performance tasks that teachers choose when to administer) can allow teachers to be more responsive to the specific learners in their classroom (e.g., if students need more time on a particular unit, want to explore a supplemental unit of study connected to that class's interests and identities, additional phenomena or alternative texts, etc). Curriculum-anchored end-of-year assessments may also draw on a narrower set of classroom work. This shift allows teachers to make more decisions about meeting the specific needs of the students in their classrooms by relying more on curriculum-embedded assessments in high-quality materials than on the state test.
- + **Decreasing interruption to teaching and learning.** More limited administrations might mean fewer interruptions to coherent teaching and learning. This effect of through-year assessments can be mitigated through the use of authentic and meaningful performance tasks (e.g., writing portfolio contributions, science investigations) that themselves provide meaningful learning experiences to students.
- + **Increasing the straightforward interpretation of assessment data.** Summative assessments that are administered once per year or grade-band make a claim about student performance at that one moment in time. While there are real limitations to this approach, through-year assessments can also pose challenges for generating and interpreting scores, as they seek to count administrations over the course of learning toward a summative assessment. Given the focus on disciplinary practices and core conceptual ideas that build over time, it is not clear how one should interpret or aggregate student performance over multiple administrations. For example, should less weight be given to assessments given earlier in the academic year?

## GOALS OF CURRICULUM-ANCHORED

- + Limiting negative impact on student experience.** Assessments are often experienced by students as traumatic, anxiety-inducing, irrelevant and disconnected from learning. More limited administrations may keep the potential negative impacts on students' experiences contained to those more sporadic assessment experiences rather than a routine part of their schooling experience. It should be noted that curriculum-anchored approaches should position state assessments to be more relevant and connected to learning; reduce some feelings of anxiety because students feel more prepared for the assessment; and, in well-designed systems, actually help students experience the assessment as compelling and motivating in its own right. However, these shifts are highly dependent on system design.
- + Being more manageable for teachers and districts.** The disciplinary considerations described above suggest that curriculum-anchored assessment systems may look very different in the different disciplines. Practically, this may be burdensome for teachers and districts as it may mean incoherent testing windows, different platforms and materials, different data systems and reporting infrastructure, etc. This may be easier to navigate in more limited administration contexts.
- + Reducing burden on the quality of the assessment.** Because through-year assessments will likely be used more directly by teachers, schools and districts to influence instruction, these assessments must necessarily be held to a higher bar of quality. More limited assessment administrations of a curriculum-anchored assessment system may focus on assessments that signal, incentivize, and exemplify high-quality teaching and learning and alignment to standards. Through-year curriculum-anchored assessments must do all of that and also provide information that provides productive and more immediate feedback for classroom activities. While it is essential that all assessments be part of continuous improvement processes to improve quality over time, the reality is that many state assessments do not yet exemplify features of high-quality teaching and learning (often for reasons outside of states' control)—given this reality, states and developers should carefully consider whether they are well-positioned to develop assessments that would meet the quality demands of a through-year assessment.

Ultimately, decisions about how states design the cadence and administration of their assessments will depend on a series of factors, including those related to feasibility and cost as well as content and use. When making these decisions, states should carefully consider:

- + Does this design attend to the most important features and shifts of teaching and learning in this discipline?**
- + Will this approach have a net positive impact on instruction for all learners, particularly those who have been marginalized in our education systems?**
- + If pursuing through-year assessments, is the design of the assessment sufficiently connected to features of high-quality curriculum?**

# CURRICULUM-ANCHORED DESIGNS

## What can we learn from current curriculum-anchored assessment systems?

Although curriculum-anchored assessment systems at the state level are a relatively novel concept in the United States, many other systems already employ assessments that are intentionally connected to curriculum (Figure 6). These existing systems provide concrete models and examples for how large-scale systems can anchor to curriculum in a variety of ways. **By considering these models, states can better understand the design considerations, affordances, and enabling conditions of different approaches to curriculum-anchored assessment systems.** These systems, described below, serve as both functional and illustrative examples of how existing and emerging assessment systems have tailored designs for discipline-specific content implications and varying timing and cadence based on their purpose.

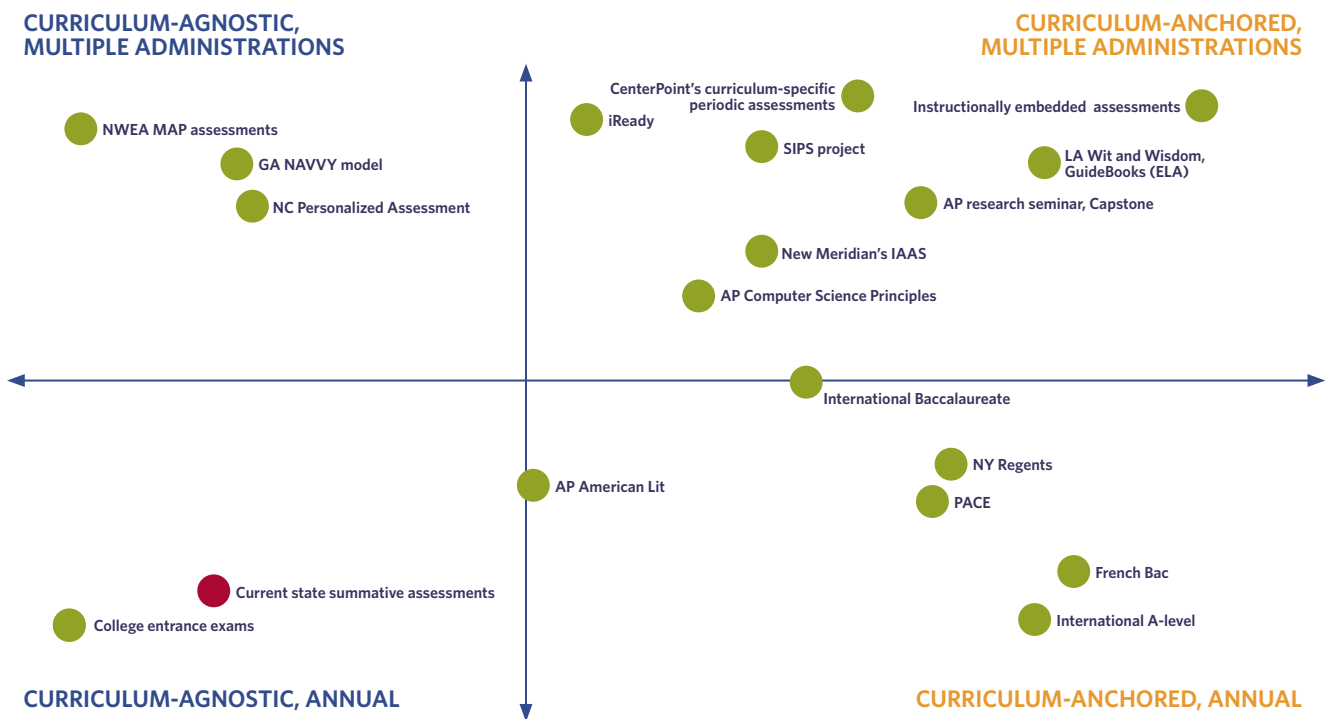


Figure 6. Examples of current assessment systems in relation to their connection to curriculum and administration cadence

### Curriculum-specific, through-year: Louisiana Innovative ELA Assessment

Each of Louisiana’s two innovative assessments in ELA are explicitly tied to a specific high-quality curriculum—currently Guidebooks and Wit & Wisdom. Sharing the same overall design, they are made up of multiple, end-of-unit assessments combined with an end-of-year writing task to produce a single summative assessment score. By being tightly coupled to the specific timing, texts and the knowledge-and-content-rich approach embedded in the instructional materials, the system is designed to assess to what degree students can (1) understand and build knowledge from the texts they have read and (2) express that knowledge and understanding in writing.

## CURRICULUM-ANCHORED DESIGNS

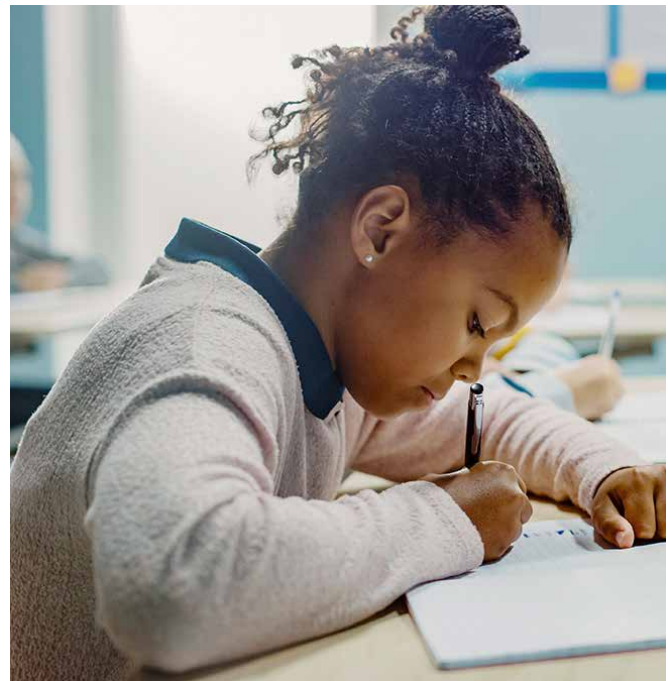
This design is meant to have four key benefits:

- + **Multiple check-ins:** Several brief assessments throughout the year, rather than one end-of-year assessment, enables the timely, in-depth review of specific materials in preparation for each periodic test administration.
- + **Instructional focus:** Teachers can focus instruction on careful, in-depth reading and knowledge building, because they know that their students will be tested on that taught material.
- + **Equity in the opportunity to learn:** All students have the opportunity to develop knowledge together through the use of rich texts so that no student is at a disadvantage due to differences in life experiences.
- + **Preservation of local control:** School systems can pick from more than one curriculum (the state is working on supporting assessments linked to further HQIM).

### Curriculum-specific, annual administration: Cambridge International A-Level English Assessment

The Language and Literature in English course provides a syllabus that specifies authors and texts to be addressed in both curriculum and assessment. On the end-of-year assessment, students are asked to consider and make sense of these specific texts with the expectation that they will have studied them during the course of instruction. There are two components to the assessment: language and literature.

- + In the language component, students build on opportunities in instruction to practice sustained, accurate, fluent and consistent writing across a variety of audiences, contexts and styles.
- + In the literature component, students address specific texts they have studied in instruction. The assessment and associated syllabus encourage students to develop their skills of analysis and interpretation as well as personal responses to the texts studied.



## CURRICULUM-ANCHORED DESIGNS

This design is meant to have the following key benefits:

- + The syllabus and text selection are intentionally designed to encourage an appreciation of literature in English—prose, poetry and drama—of different types and from different cultures.
- + By specifying a range of texts, students are able to develop the key knowledge and skills required to read, analyze and communicate effectively in English in more generalizable ways.
- + Learners develop deep understandings of the subjects they engage with in the texts, understand more about writers’ choices of language, form and structure and develop their ability to form independent opinions about what they read.
- + Learners also improve their understanding of the English language and how it is used by extending their skills across a range of writing styles, including imaginative, discursive and argumentative.
- + The list of texts is fluid which enables the assessment designers to reflect current cultural expectations while also preserving a place for a varied list of canonical works.

### **Instructional model-connected, embedded tasks + sit down assessment: International Baccalaureate**

In the International Baccalaureate’s (IB) primary years, middle years and diploma programmes (PYP, MYP and DP respectively), students are assessed using a combination of sit-down, end-of-year tests and performance assessments (e.g., projects, independent research, laboratory activities, etc) that include both common tasks as well as locally developed and scored papers and projects with external centralized moderation. While educators and students often have some flexibility in choosing the specific components of instruction, the assessment system is closely connected to both the IB’s overarching instructional model and learner profile (e.g., focusing on global and intercultural considerations, developing theories of knowledge, interdisciplinary activities, etc.) as well as to discipline specific considerations. For example, see Table 2 for some curriculum-anchored features included in assessments diploma programmes.



# CURRICULUM-ANCHORED DESIGNS

**Table 2: Examples of curriculum-anchored features found on assessments for International Baccalaureate's diploma programmes**

	LITERATURE	BIOLOGY	MATHEMATICS
Assessment connections to IB instructional model	Examination of how texts in their language of origin as well as a translated text have approached a common global issue	Student developed individual investigation and interdisciplinary and collaborative science project	Mathematical exploration that focuses on student demonstration of knowledge and skill in an area of personal interest
Discipline-specific assessment-curriculum connections	Includes both previously studied and related but unseen texts	Focuses on evidence-based analysis and scientific reasoning using the major science ideas and practices developed in instruction for the majority of the sit-down assessment.	Focuses on problem-solving in non-routine, open-ended and real-world contexts, and focuses on the development of conceptual understanding in addition to procedural skills
Relevant notes	The assessment leverages both some common texts from instruction as well as elements of the instructional model. This allows for local choice in terms of some texts, while ensuring that the assessment is both driving and measuring important instructional activities.	By including authentic individual and collaborative projects in addition to a sit-down assessment that focuses on key behaviors of the discipline, the assessment is coherent with the expected instruction.	The assessment matches key distinguishing features of the curriculum.

The IB examples are notable because IB assessments must provide scores used as part of high-stakes decisions (e.g., graduation status, college credit) while operating within an extremely wide range of cultural and policy contexts, including over 150 countries and the vast majority of states in the United States. By focusing on alignment to the instructional model rather than to specific materials, IB assessments have a strong reciprocal connection with teaching and learning while allowing considerable flexibility in terms of specific curricular implementation.

## Curriculum-embedded extended performance task: AP Computer Science Principles

The AP Computer Science Principles exam includes both a standard, on-demand assessment that is aligned to the learning goals of the course as well as a highly authentic, curriculum-embedded performance task where students collaboratively design, test and describe a computer program. By including the performance task as well as the expectation that class-time will be used to do the task, the assessment is functionally setting part of the curriculum. This aspect of the assessment design positions the assessment to be both particularly instructionally relevant and educative to

## CURRICULUM-ANCHORED DESIGNS

both teachers and students; it ensures that assessment and curriculum are tightly coupled to each other. At the same time, teachers have a great deal of flexibility to choose programming languages, surrounding teaching and learning context, etc.

### **Scope and sequence aligned, multiple administration through-year: New Meridian Instructionally Aligned Assessment System approach in mathematics and ELA**

New Meridian is developing a testlet approach with states that seeks to carefully walk the line between focusing on standards and providing instructionally relevant information that is connected to the curriculum students are experiencing. In this model, a through-year assessment is made up of micro-assessments that districts and schools can select and use when appropriate to teaching and learning in their contexts. These micro-assessments focus on highly specific attributes of student performance that are linked to specific teaching and learning elements associated with building toward the grade-level standards. For example, items measuring the Common Core math standards related to performing operations with decimals may be designed to distinguish between students who can subtract fluently when the decimal points are in the same position in both terms but who cannot subtract fluently in other instances. New Meridian is currently working on ways to aggregate performance on the testlets to provide valid and reliable measures of how well students have mastered grade-level standards. By providing information along the developmental progression toward grade-level standards at appropriate times in instruction, the testlets can be used to provide information about student performance relative to grade-level standards while also being connected to the high-quality curriculum being used.

### **Scope and sequence aligned, curriculum-specific, multiple administrations throughout the year: CenterPoint’s interim approach in mathematics, ELA, and science**

CenterPoint has been developing interim assessments in math, ELA, and science (forthcoming) that intentionally align with standards as well as both the scope and sequence as well as other distinctive design elements—including developmental progressions—specific to each set of instructional materials. In this model, CenterPoint designs the interim assessments in direct collaboration with high-quality instructional materials developers. This process produces assessments that students and teachers experience as coherent with, and part of, teaching and learning, while simultaneously providing a common, external validation of student progress that can be used by classroom, building and district leaders as well as parents and families. By working in close collaboration with instructional materials writers, CenterPoint’s assessments leverage the careful work underlying the instructional models, developmental progressions, and specific contexts, texts, content and scenarios present in the curriculum, allowing the assessments to pay careful attention to both opportunity to learn up to the point of the assessment as well as next steps teachers may take to better support learning as the curriculum continues to unfold.



## FEDERAL AND LOCAL NEEDS



### How can states design curriculum-anchored assessment systems that balance federal requirements for state assessments and local curriculum decisions?

**M**ultiple curricula are being used within every discipline, at every grade-level, in every state. Decisions about curriculum, instructional materials and professional learning are largely local decisions, and no state mandates the use of a specific curriculum. Moreover, many teachers, schools and districts adapt curricula in various ways and to various degrees. Despite all these factors, states need to ensure that any assessment given to students can generate comparable results such that teachers, schools and districts can ensure all students are supported in creating strong learning outcomes. How can a state develop curriculum-anchored assessment systems that meet both of these needs?

Curriculum-anchored assessment systems will need to carefully attend to this balance. While there may not be a perfect solution, there are a range of ways states can begin considering these systems.

#### How can states attend to a diverse and evolving curriculum landscape?

There are several ways curriculum-anchored assessment systems can allow for the use of multiple high-quality curricular choices. These options can largely be organized into two major categories:

- + Assessment designs that **respond to** high-quality curriculum. In other words, assessments that are designed based on existing HQIM. In these systems, the content of the instructional materials has a direct (and somewhat unidirectional) influence on the assessment design.
- + Assessment designs that **set curriculum**. In these designs, features of the assessment are used to intentionally drive some curricular choices (e.g., by clearly stating that certain texts, phenomena, investigations, etc will be assessed). States may design these elements in ways that connect and incentivize the use of some existing HQIM, or these elements may be designed in ways that shape the design, adaptation, adoption, and implementation of instructional materials to come.

## FEDERAL AND LOCAL NEEDS

States could explore variations on each of these options, depending on the right trade-offs for that state. Importantly, assessment designs that “respond to” and “set” curriculum can both be accomplished across a spectrum of curriculum anchors, utilizing tighter or looser couplings to specific curriculum.

**Table 3. System Design Options that Attend to Multiple Curricula**

	OPTION	RELEVANT EXAMPLES
Respond to high-quality curriculum	<b>Multiple curriculum-specific assessments.</b> The state develops and allows multiple assessments. Each assessment is both aligned to state standards and designed to specifically be used with a given HQIM.	Louisiana Innovative Assessment
	<b>Curriculum-specific elements of assessments coupled with a common test.</b> The summative score is based on two elements: One component of the state assessment is responsive to curriculum-specific materials (e.g., questions enabling students to show curriculum-related knowledge and/or use of curriculum-embedded performance tasks) while a second component connects to common features or topics found across relevant curricula (i.e., individualism in American Literature).	AP American Literature, International Baccalaureate.
	<b>Align to instructional models, shifts and features shared by multiple high-quality curricula.</b> The state can look to discerning common features shared across high-quality curricula and choose to center those features in the assessment design (e.g., research-based developmental approaches, common topics).	International Baccalaureate, New Meridian instructionally aligned assessment system
Set high-quality curriculum	<b>Make public certain components of the assessment.</b> The state makes clear certain texts, topics and task types that will be assessed and effectively uses the assessment to set some core curriculum content.	British A-levels
	<b>Provide curriculum components that will be incorporated as part of the summative.</b> The state can provide performance assessments, exemplar units, investigations, etc that are locally administered and either will be scored as part of the summative or will be included in some other way (e.g., students will be asked about the texts, investigations, etc, on the summative).	AP Computer Science Principles

### How can states design curriculum-anchored assessment systems that meet federal requirements for state assessment systems?

State assessments in mathematics, ELA and literacy and science must meet federal requirements for assessments. Such requirements currently include that all students in a state are given the same assessment, that assessments generate comparable student scores and that assessments align to the state’s standards. Some design options may pose challenges for meeting these requirements, as they depend on how assessments are anchored to the curriculum. However, there is a path forward in nearly all cases. Some options are very straightforward; others involve more intentional (and potentially creative) approaches to alignment and psychometrics while still operating within the

## FEDERAL AND LOCAL NEEDS

spirit of guardrails designed to ensure all students are assessed fairly and supported in meeting high standards; and still other options may involve agreements with the U.S. Department of Education that may lead to potential waivers, additional innovation authorities or other shifts in federal policy.<sup>13</sup>

Below, we discuss (1) some of the more challenging technical issues states will need to address when designing assessments to meet federal requirements and (2) policy pathways that may support states in doing so.

### Technical Issues: Options for States to Consider

There are at least three major technical issues related to compliance with federal requirements that states encounter when considering curriculum-anchored assessment systems. These include:

1. Ensuring that assessments are aligned to the depth and breadth of a state’s grade-level standards and that assessments will provide scores that accurately describe achievement relative to those targets.
2. Ensuring that all students are given the same assessment, containing comparable tasks that generate comparable results.
3. Ensuring that assessments generate valid, reliable scores every year by maintaining expectations for test-security.

Below, we discuss ways states may explore addressing these concerns. It should be noted that while the Every Student Succeeds Act (ESSA) allows quite a bit of flexibility regarding how students are assessed—and explicitly encourages more innovative and instructionally relevant assessment approaches—both the regulations and the processes for approving state assessments (e.g., federal peer review) established by the U.S. Department of Education have largely been developed for current curriculum-agnostic end-of-year assessments. This is also true of many of the partners and systems working with states; the field has been shaped by the current status quo. Many of the processes, requirements and trade-offs made in state assessments privilege current approaches, even if curriculum-anchored assessments may offer higher quality assessments and information. This means that designing assessments that truly connect meaningfully with curriculum will likely require that states, in partnership with external experts and the U.S. Department of Education, carefully consider (1) assessment purpose, (2) the appropriate trade-offs to make and (3) where to advocate for judicious deviations from current practice versus where to design within current criteria and guardrails. States will need to have open and ongoing conversations with the U.S. Department of Education to ensure that their designs can move forward. Table 4 describes considerations states should attend to in particular for different curriculum-anchored options.

<sup>13</sup> For an example of precedent for flexibility, see the use of a long-standing alternative assessment for several districts in New York State. See Hoff, D.J., (2005).

# FEDERAL AND LOCAL NEEDS

*Table 4. System Design Options that Attend to Multiple Curricula: Connection to Federal Requirements*

	OPTION	WHICH CURRENT FEDERAL REQUIREMENTS POSE THE MOST SIGNIFICANT BARRIERS?	PATH FORWARD
Respond to high-quality curriculum	Multiple curriculum-specific assessments.	Test security, comparability, requirement for the same test	Discussions with USED and Congress, explore changes in law with appropriate guardrails (e.g., new innovation opportunities, changes to ESSA section 1111 over time)
	Curriculum-specific elements of assessments coupled with a common test.	Test security, comparability, requirement for the same assessment	Discussions with USED, strategic demonstration of evidence to meet peer review criteria
	Aligning to instructional models, shifts, features shared by multiple high-quality curricula.	Alignment	Strategic demonstration of evidence to meet peer review criteria
Set high-quality curriculum	Make public certain components of the assessment.	Test security	Discussions with USED, strategic demonstration of evidence to meet peer review criteria
	Provide curriculum components that will be incorporated as part of the summative.	Test security	Discussions with USED (note that this could be more or less challenging to include depending on how the curriculum component will be included in the summative assessment)

Key (given current federal law and requirements):  **Challenging**      **Moderate**      **Easiest**

## Alignment to the depth and breadth of grade-level standards

A major value-add of curriculum-anchored assessment systems is that it can help states operationalize a high-quality, instructionally-useful interpretation of standards. High-quality curricula, including instructional materials, are most frequently the result of:

- + Hundreds of hours of standards unpacking,
- + Intentional collaboration between diverse educators, content experts and learning scientists specializing in students’ disciplinary learning,
- + Feedback and design support from students, and
- + Careful consideration of not only what students should know by the end of instruction (defined by standards) but also how students should get there, given a realistic understanding of their starting points and an understanding of how students learn.

## FEDERAL AND LOCAL NEEDS

In short: high-quality curriculum translates standards from abstract skills into concrete understanding and practice—and as a result, offers a better resource for designing assessments aligned to standards than the standards alone. However, aligning to standards via curriculum requires that states seek out intentional approaches to documenting evidence of alignment that accounts for the specific ways curriculum-anchored assessments may interpret standards and assess student achievement.<sup>14</sup>

For example, many alignment methodologies focus on superficially matching items to standards in ways that might seem reasonable for a large-scale, standardized assessment, rather than carefully assessing whether the items and tasks used on an assessment are actually surfacing evidence of student progress toward and performance of standards. In curriculum-anchored systems, test items will—for good reason—likely be designed differently. Items and tasks may be designed to assess learning along progressions toward grade-level standards to help educators support sequential, cumulative student learning. In some cases, an item may only assess part of a standard because the curriculum separates how students learn that standard across multiple units. In other instances, items and tasks may bundle multiple standards together.

In all of the instances above, the resulting test may be more supportive and reflective of students' opportunity to learn, more useful to instruction and in fact yield results that more authentically represent student achievement relative to the depth and breadth of grade-level standards than the pre-existing state tests. Unfortunately, a cursory look at the test could lead to judging it as misaligned under current approaches and frameworks.

When considering appropriate approaches to alignment for curriculum-anchored systems, states and partners may consider pursuing **clear alignment methodologies that account for intended purposes, curriculum-based considerations and overall integrity of state standards across each academic year**. This may involve leveraging panels of experts and educators, clarifying the role of HQIM evaluations in determining assessment alignment and deciding the appropriate range of evidence needed for curriculum-anchored assessments designed for different purposes. These steps will require that federal peer review processes are also updated to accept these kinds of alignment studies. There is good reason to believe this is achievable, based on prior examples of department approval for state assessments that used different, updated approaches to demonstrating alignment to state standards.

---

<sup>14</sup> Note that a more nuanced interpretation of standards may also have implications for more useful achievement level descriptors/achievement standards created as part of assessment development processes. For states pursuing curriculum-anchored assessment systems to simply replace their current assessment, this will not pose an issue. However, states seeking to pursue curriculum-anchored assessment systems under IADA are required to show comparable scores relative to the achievement standards of the current assessment—which poses some additional challenges for states seeking to fully redesign their system.

## Comparable scores

Ensuring that assessments produce valid, reliable and comparable scores is essential for states to ensure that they can surface and support needs within schools, districts and communities. As states consider curriculum-anchored assessment systems, they will need to consider how to provide assessments that can intentionally support and reflect multiple curricula. Some states will consider multiple assessments, each designed for/aligned to a specific curriculum. In states seeking multiple, curriculum-specific assessments, this issue is twofold: (1) ESSA requires that all students be given the same assessment (with some notable exceptions<sup>15</sup>) and (2) states need to demonstrate that the assessments in their statewide system are comparable instruments that generate comparable results among all students taking the curriculum-specific assessments. Additionally, under IADA, if a new test design creates a different pattern of results than the existing test (even if this is the result of intentional design to assess components of the standards that are underrepresented or ignored in the current test), that difference positions the new assessment as out of compliance with the expectations of the demonstration authority.

How can states demonstrate comparable scores across students, schools and districts and simultaneously attend to multiple curricula? Answering this question requires that states be very clear: **Comparable to what** and **comparable for what purpose**?

Depending on the answers to those questions, there are at least two options states can consider. One option is to work with the Department of Education and assessment experts to identify **approaches to demonstrate comparability that require comparable tasks and expectations instead of identical tests**. Although comparability is often interpreted as expectations for identical tasks and tests, this is an outdated idea and not a requirement for data that can be used to make trustworthy judgements about comparisons of student performance. Indeed, by ensuring that assessments are flexible enough to provide students with fair opportunities to demonstrate their learning—such as by attending the topics, texts and experiences they had the opportunity to learn—assessments may be able to generate even more trustworthy comparisons than current, curriculum-agnostic assessments. For example, many high-stakes literature assessments (e.g., AP American Literature, IB English) allow curriculum-specificity by asking the same questions of students but enable them flexibility in the choice of texts and/or specific topic content in their answers. These assessments are effectively asking students to still take the same test, because the task demands (and the standards students will be evaluated against) are the same across all

---

<sup>15</sup> There are some notable exceptions to this requirement that suggests there is opportunity for flexibility given conversation with USED and appropriate justification. For example, the KA'EO assessment program in Hawai'i is an approved state assessment that provides students who participate in the Hawaiian language immersion schools (Kaiapuni Schools) with a culturally appropriate Hawaiian language assessment in lieu of Smarter Balanced math and ELA assessments and the state's science assessment. Similarly, the 38 high schools in the New York Performance Standards Consortium received a waiver from New York State from participating in the state's Regents Exams in all content areas except ELA (which is still required for NYPSC students to receive a Regents diploma in the state). In both Hawai'i and New York, the assessment program was able to demonstrate their assessments were of equal or better quality than the traditional statewide summative assessment. These examples illustrate mechanisms, including approval from USED and internal state waivers, for states to support multiple assessment systems when appropriate.

## FEDERAL AND LOCAL NEEDS

students, regardless of curriculum. It is not difficult to imagine a similar model being applied to state assessments that would allow students to use texts, topics, phenomena and problems that are specific to their learning experiences but within common tasks and relative to common standards.

It should be noted that this approach does require reporting against grade-level standards but does not preclude providing additional information relative to curriculum-based learning goals, attributes and other features of student performance that might be particularly useful in teaching and learning contexts, if the assessment was designed to provide this level of information.

States could also consider including **substantial common elements—in terms of timing, administration and content—across assessments designed for multiple curricula to demonstrate comparability**. This is functionally the path many programs and states pursuing innovative assessments are exploring, including:

- + States considering through-year assessments that only use the final administration for summative scores,<sup>16</sup>
- + Those IADA<sup>17</sup> states that are designing new assessments with both an innovative component and a component of their current state assessment (with substantial overlap between the innovative assessment and the current statewide assessment),
- + States considering using thematic and design bridges, rather than scoring inclusion, between the state summative assessment and curriculum-based assessment elements (e.g., performance tasks) to more meaningfully connect with teaching and learning, and
- + Programs like AP Computer Science that include authentic curriculum-embedded performance tasks that are scored and used together with an end-of-year, on-demand assessment to generate students' final scores.

In a curriculum-anchored model, this could look like (1) a common on-demand component of the assessment administered to all students, regardless of the curriculum model that governs the rest of their assessment experiences, (2) common tasks administered within the curriculum-anchored models that are designed to be comparable enough to provide the needed scores or (3) a combination of both approaches.

<sup>16</sup> See Education First (2022) for further detail on how different states are considering calculating summative scores.

<sup>17</sup> Currently, IADA requires that states demonstrate that their innovative test generates comparable results/scores to the existing assessment. The Every Student Succeeds Act requires states to demonstrate that the innovative assessment system will “generate results that are valid and reliable, and comparable, for all students and for each subgroup of students described in section 1111(b)(2)(B)(xi), as compared to the results for such students on the State assessments under section 1111(b)(2).” See ESSA, Section 1204(e)(2)(A)(iv).

## FEDERAL AND LOCAL NEEDS

### Test Security

States are currently required to adhere to strict test security requirements in an effort to make sure scores represent what students actually know and can do and do not unfairly benefit from practices such as premature test question exposure. While well-intentioned, this may pose challenges for curriculum-anchored assessment systems. For example, curriculum-anchored ELA assessments may seek to leverage known topics and texts and even known tasks (e.g., an expectation for synthesis)—and, in many instances, the same fixed curriculum will be used year after year. In this latter case, the assessment would have to keep generating new questions on the same material or take hard-to-imagine steps to keep prior questions secret.

There are two sides to this challenge: designing curriculum-anchored assessments that meet the spirit of the test security requirements to ensure valid scores and actually meeting the letter of the requirement for state assessments laid out in **federal peer review guidelines**. There are a few ways states can approach the former, depending on their test design and intended outcomes:

- + Focus a substantial amount of the assessment on authentic tasks and performances connected to the curriculum.** Perhaps one of the most compelling ways to address the underlying rationale for test security is to ensure that major components of the test require authentic performances (e.g., synthesis of texts, investigation design), such that prior knowledge of the task and related materials do not compromise validity of scores—and may even make the scores more valid, as described above. This is particularly obvious in ELA, where both appropriate background knowledge and prior experience with some texts actually creates opportunities for deeper assessment of what students know and can do, and similar approaches make sense in other disciplines as well. For example, in AP Computer Science Principles, students are asked to design a computer program and submit both written and oral descriptions of code, how the app works, etc. Knowing this is a requirement beforehand does not invalidate the performance—and, in fact, this kind of coupling between curriculum and assessment is a goal of many curriculum-anchored systems. The key lies in tasks that are authentic representations of what and how students should be learning.
- + Determine appropriate ways to refresh texts, related phenomena, source materials, etc.** Another path toward test security may be to consider how frequently some static elements of curriculum-anchored assessments (such as texts, phenomena, source materials, etc) can be refreshed. In curriculum-anchored assessment systems that rely on connections to curriculum that is not changing, it may make sense to (1) focus refreshing elements of the assessment that are intentionally targeting transfer of understanding (e.g., warm and cold reads in ELA, phenomena in science) in each of the disciplines and/or (2) rotate which elements of the curriculum (e.g., which texts, phenomena) are used in the summative design.
- + Take different, appropriate approaches to test security for different assessment components.** As states design curriculum-anchored assessments, they may find that their assessment designs include diverse enough components that different approaches to test security can be used for different pieces. For example, states may have some largely selected response sections



## FEDERAL AND LOCAL NEEDS

of the assessment for which it makes sense to follow traditional approaches to security, while others that use performance tasks (which need not be kept confidential). Some disciplines may lend themselves to large pools of items that could all be released, such that memorizing responses to all of them is unreasonable.

Addressing the other component of the test security requirement (i.e., demonstrating evidence of meeting the requirement for federal peer review) may require follow up with the U.S. Department of Education (discussed further below).

### Policy Pathways

Curriculum-anchored assessment systems are innovative in the United States and may require some support to implement. In addition to simply replacing a state's current assessment system, which states can do at any time, there are at least three policy pathways states can pursue to more gradually pilot and enact curriculum-anchored assessment systems that meet federal requirements. These may be pursued in parallel with actions the Department of Education may take to support these kinds of innovative, instructionally-relevant assessment systems.<sup>18</sup>

- + Leverage existing opportunities for innovation, such as the Innovative Assessment Demonstration Authority (IADA) and the Competitive Grants for State Assessments (CGSA).** A limited number of states can use IADA as a way to establish curriculum-anchored systems such as Louisiana has done. This requires that states have already designed and potentially piloted their curriculum-anchored systems. It is important that states understand the full set of expectations of IADA, including potential barriers, before committing to this path. For many states, it will be more straightforward to design a new system outside of the bounds of IADA. For all states considering a new system, the CGSA program offers one path for both conversations with USED and funding in order to explore curriculum-anchored assessment systems. In recent competitions, USED awarded up to \$3M to states and groups of states to pursue assessment innovations and activities.
- + Pursue curriculum-anchored assessment activities outside of accountability-connected testing.** States may consider pursuing curriculum-anchored elements of assessments in parallel to ongoing accountability testing. While the exact approach will vary in states, states may consider support for curriculum-aligned interim assessments, access to high-quality curriculum-connected performance assessments and support for better integrating assessment systems embedded within HQIM into states' comprehensive assessment systems.
- + Request support (potentially including waivers) from the U.S. Department of Education.** States interested in fundamentally new systems may also work directly with the U.S. Department of Education to figure out the right policy avenues and flexibilities to make their visions a reality. This could lead to state-specific waivers, updates to existing innovation pathways (e.g., IADA, CGSA allocation) or potentially to new pathways for innovation all together.

<sup>18</sup> For a more extensive discussion, see Badrinarayana, A., & Darling Hammond, L. (2023) and Education First (2023).

## ENABLING CONDITIONS



### What are the enabling conditions in states?

**A**s states consider creating curriculum-anchored assessment systems, certain conditions may make these systems more feasible. Some of these conditions involve enabling federal policies and supports (as discussed above), but many have to do with states' internal contexts. These enabling conditions include:

- + **Coherent vision for assessment as part of teaching and learning.** Curriculum-anchored assessment systems fundamentally reposition state assessments as part of teacher, school and district teaching and learning strategies. This process requires that states have a clear vision for how their assessments will serve effective teaching and learning, what unintended consequences educators and students may encounter and how to monitor progress and navigate any challenges that arise. This vision and plan should put the impact on students and teachers as the most important consideration for assessment system design.
- + **Clear state-wide instructional materials strategy and data.** To design curriculum-anchored assessment systems tied to HQIM, states will need to have a clear instructional materials strategy. This should include ongoing processes for determining which materials are considered high-quality; defining the state role in establishing 'approved' materials lists and incentivizing strong curricula choices; tracking the adoption and uptake of HQIM; and monitoring student and teacher experience with the adopted HQIM.
- + **Culture of systemic support for teaching and learning.** One of the most important enabling conditions for statewide curriculum-anchored assessment systems is the buy-in of stakeholders, including educators, students and families. Their buy-in depends, in part, on the culture and history of education systems in the state, including:
  - + SEA-community relationships and whether the SEA has traditionally been seen by districts and teachers as a positive partner in teaching and learning,
  - + Whether educators and leaders see centralized curriculum-related policies that incentivize the use of externally-developed instructional materials as an important tool for supporting learning rather than a constraint on teachers' professional judgment, and
  - + Ongoing cross-district collaboration within the state.

## ENABLING CONDITIONS

- + **Consistent leadership.** Curriculum-anchored assessment systems require substantial and sustained resources and commitment as well as a clear understanding of a state’s current and evolving context. Consistent leadership at the SEA, including both policy and content leaders, is essential if curriculum-anchored systems are to take root and flourish. In a time of particularly high turnover of personnel within SEAs, states should consider distributed leadership of such initiatives as well as the role of state partners (e.g., universities, intermediary organizations) in providing some continuity over time.
- + **Dedicated time and financial resources.** Like all innovative projects worth pursuing, new assessment systems require the time and support to develop, pilot and iterate, and to allow for failing forward as needed. States should consider how they can leverage internal projects and resources as well as partnerships with other states, foundations and partner organizations to create the needed time, space and capacity.
- + **Emerging evidence and pilots.** As more states and systems pursue these systems, the field will all become smarter about how to establish effective curriculum-anchored systems. Emerging evidence and pilots will provide states with lessons learned and compelling strategies they can use to make informed, effective decisions about curriculum-anchored systems in their own states.



## CONCLUSION



## Conclusion

Curriculum-anchored assessment systems offer states a way to support teaching and learning while monitoring and supporting progress across their schools and districts. There are a wide range of ways states can consider anchoring their assessments in high-quality curriculum, but some of the most important considerations states will need to consider are disciplinary implications for assessment design, timing and cadence of assessment administration, and capacity and conditions for the implementing these systems successfully within diverse state contexts.

While curriculum-anchored assessment systems represent an evolution of state assessment systems that better reflect both the science of learning as well as lessons learned from current assessment systems, there are still several questions that are worth considering as these systems emerge.

- 1. How can curriculum-anchored assessment systems attend to culturally responsive, relevant and sustaining pedagogies?** Perhaps one of the most significant tensions underlying the development of curriculum-anchored assessment systems is the potential to constrain what happens in teaching and learning. Under some circumstances—such as motivating a shift away from low-quality instructional materials—this ‘constraining’ feature is seen as a positive aspect of curriculum-anchored systems. However, these same systems should not constrain teachers and schools as they work to be responsive to dynamic student populations (e.g., designing lessons to build on the specific cultural and linguistic assets of either the local community or a particular classroom of students, choosing texts and phenomena that connect with the learners in the room).

Given the decades of research on how vital students’ cultural and linguistic funds of knowledge are to learning, it will be imperative to ensure that these elements are not inadvertently lost in an effort to connect curriculum and assessment more directly. Future work may consider (1) whether some system designs are more conducive to culturally relevant and sustaining teaching and (2) whether there are features of instructional materials that can be included as part of quality reviews to ensure that assessments are reinforcing HQIM content that attends to culture in meaningful ways.<sup>19</sup>

<sup>19</sup> It should be noted that current state assessments are egregious on this point, and curriculum anchored assessments at the very least attend to shared experiences students have in the classroom as one element of their funds of knowledge to build upon.

## CONCLUSION

- 2. How can curriculum-anchored assessment systems avoid constraining needed evolution of instructional materials?** Generally speaking, HQIM is not designed to be a completely static resource once developed and released—indeed, most HQIM developers are in a constant state of field-testing, iterating and diversifying their materials. This is quite foreign to many state assessment development processes: mechanisms which can be so burdensome that any significant deviations in test design are deeply discouraged, if not labeled impossible. Additional work will need to be done to develop solutions for state assessments that do not constrain the necessary evolution and improvement of instructional materials used to ground curriculum-anchored assessments.

One interesting option states may consider is whether instructional materials developers may themselves become partners in the state assessment design. Many HQIM already have robust and meaningful assessment systems embedded within their materials and are explicitly designed to both monitor student progress toward mastery and provide information to help teachers and students implement the materials more effectively. It is conceivable that, over time, either parts of these very same embedded assessment systems could become part of the state system or that instructional materials developers could develop separate but coherent assessments that grow with the curriculum more organically.

- 3. How can curriculum-anchored assessment systems be educative to novice and master teachers alike?** HQIM play a different role for teachers who are still learning how to effectively teach within a discipline (and may be more likely to benefit from teaching the materials with fidelity) than for those who are adept at their craft and may be able to teach to the spirit of the materials but with appropriate adaptations for their teaching context. It will be imperative that states considering curriculum-anchored assessment systems consider their current and expected teacher workforce and the state of current teaching practice in each of the disciplines to determine how assessments can be designed to support teachers at very different levels of pedagogical experience and expertise.
- 4. Given the increasing focus on relevant multi- and interdisciplinary teaching and learning efforts, how should curriculum-anchored assessments attend to multiple disciplines when needed?** The question of multi- and interdisciplinary assessments has long been a topic of interest. This is especially true as the disciplinary standards have come to more closely reflect how math, ELA, science and social studies are practiced in the real world—not as siloed as course titles or assessment scores may lead one to believe. Future work should explore models that allow for appropriate interactions among the disciplines while still accounting for most schools’ disciplinary-based approach to courses and teaching.
- 5. What assessment systems can be used to validate HQIM?** One common pushback against curriculum-anchored assessments is that the assessments should be able to play a role in validating the use of a particular curriculum—and that by connecting the assessment to a particular curriculum, this process cannot occur. One response is that nearly every curriculum-anchored assessment system includes intentional assessment of transfer relative to the learning context and alignments to standards. Nevertheless, future research should focus on ways to validate HQIM, including through student work and performance on aligned assessments as well as appropriate options for external assessments.

## CONCLUSION

### Looking Ahead

Over the course of the 2022-23 school year, the through-year curriculum-connected assessment grant program has seeded multiple new assessment designs and test prototypes. It is too early to tell whether these innovations are delivering on the marked improvements that would justify their development into full-scale operational systems. However, these research and development efforts are producing valuable lessons for the field.

At the onset of this project, we developed a research agenda informed by the piloting states and assessment developers as well as an advisory group of State Education Agency (SEA) leaders. We aimed to illuminate how states are addressing practical challenges and weighing trade-offs as they pilot their through-year assessment models. In consultation with state leaders and assessment experts, we've identified core questions at different stages of the research and development process:

1. How should states determine if a through-year assessment system aligns with their state goals?
2. What infrastructure and enabling conditions must be in place to implement a through-year assessment system and support transitioning to a new assessment system?
3. How can through-year assessment designs move toward coherence with what is being taught when many curricula are used within a state?
4. How are states addressing the logistical and operational implications of having multiple state assessment administrations?

We have produced a **primer on through-year assessments** being developed across the country: a document that surveys the design choices states and their assessment partners are making (both those intentionally connecting to curriculum and those that are not). We have also produced a **toolkit** which aims to help state leaders wrestle with the first two core questions above.

Our next paper in the series will address the final core question above: "How are states addressing the logistical and operational implications of having multiple state assessments instead of a single summative?" Drawing from lessons learned by states implementing through-year assessments, this report will provide illustrative examples of how states are addressing the implementation challenges of conducting multiple summative administrations throughout the year (e.g., cost, test security, student mobility) and are actualizing the promise of through-year assessments to provide better feedback to teachers, students and families (e.g., results reporting, building assessment literacy, professional learning). We aim to show how these challenges can be addressed without elevating logistical burdens, instructional interruptions and pressure on students and educators.

This series of papers and tools can be found on a **microsite** devoted to this project, and we hope you will join us in wrestling with how these innovations can inform how we create state summative assessments that are more equitable, relevant and coherent for students, families and educators.

## AUTHORS & ACKNOWLEDGEMENTS

### Authors



Aneisha Badrinarayan

**Aneisha Badrinarayan** is Director of State Performance Assessment Initiatives at the Learning Policy Institute, where she supports states in developing meaningful, instructionally-relevant systems of assessment. She previously served as the Director for Special Initiatives at Achieve, museum professional and a researcher focused on learning and memory encoding in the brain.



Dr. David Steiner

**Dr. David Steiner** is Executive Director of the Johns Hopkins Institute for Education Policy and Professor of Education at Johns Hopkins University. He previously served as the NYS Commissioner of Education, the Klara and Larry Silverstein Dean of the School of Education at Hunter College, CUNY, and the Director of Arts Education at the National Endowment for the Arts. He is the author, most recently, of *A Nation at Thought: Restoring Wisdom in America's Schools*.

### Acknowledgements

We are deeply grateful to all of the state leaders, curriculum and assessment experts, practitioners and policy leaders whose ideas, victories and challenges have influenced the thinking laid out in this paper. While the insights here—including any error of fact or by omission—are ours alone, notions of curriculum-anchored assessment systems are evolving through years of work by many individuals and organizations. In particular, we thank the following people and teams for generously sharing their time and insight on the relationship between curriculum and assessment systems:

- + **Johanna Brown**, Washington Office of Public Instruction
- + **Sara Cooper**, CU Boulder
- + **Ted Coe**, NWEA
- + **Michael Cohen**, CenterPoint Education Solutions
- + **Nathan Dadey**, Center for Assessment
- + **Miah Daughtery**, NWEA
- + **Ellen Ebert**, Washington Office of Public Instruction
- + **Erin Furtak**, CU Boulder
- + **TJ Heck**, CenterPoint
- + **Alissa Kilpatrick**, Independent Consultant

## AUTHORS & ACKNOWLEDGEMENTS

- + **Tana Luther**, Louisiana Department of Education
- + **Scott Marion**, Center for Assessment
- + **Sharon Necaise**, Louisiana Department of Education
- + **Jim Pellegrino**, University of Illinois, Chicago
- + **William Penuel**, CU Boulder
- + **Breigh Rhodes**, Louisiana Department of Education
- + **Sam Ribnick**, Massachusetts Department of Elementary and Secondary Education
- + **Laura Slover**, CenterPoint Education Solutions
- + **Corrine Steever**, Massachusetts Department of Elementary and Secondary Education
- + **Molly Talbot**, Louisiana Department of Education
- + **Arthur VanderVeen**, New Meridian

**The Through-Year Curriculum-Connected Assessment Grant Program founded in 2021 by the Bill & Melinda Gates Foundation, the Walton Family Foundation, the Chan Zuckerberg Initiative and Education First, was developed to support this work. To learn more about this work, please visit:**

**<https://www.education-first.com/innovations-in-assessment/our-impact-through-year-curriculum-connected-assessments>**



Badrinarayan, A., & Dadey, N. (in prep). Reimagining state assessment as a positive force for instruction.

Badrinarayan, A., & Darling-Hammond, L. (with DiNapoli, M., Kini, T., Miller, T., & Woods, J.). (2023). Developing state assessment systems that support teaching and learning: What can the federal government do? Learning Policy Institute.  
<https://doi.org/10.54300/885.821>

Cambridge exams. (n.d.). Cambridge International AS and A Level English Language and Literature (AS Level only) (8695). Retrieved from  
<https://www.cambridgeinternational.org/programmes-and-qualifications/cambridge-international-as-and-a-level-english-language-and-literature-as-level-only-8695/>

CCSSO Instructional Materials and Professional Development Network. (n.d.). High-Quality Instructional Materials. Council of Chief State School Officers. Retrieved from  
<https://learning.ccsso.org/high-quality-instructional-materials>

Dadey, N., and Badrinarayan, A. (2022). In search of the just-right connection between curriculum and assessment. The National Center for the Improvement of Educational Assessment. Retrieved from  
<https://www.nciea.org/blog/in-search-of-the-just-right-connection-between-curriculum-and-assessment>

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

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>

Doorey, N., & Polikoff, M. (2016). Evaluating the content and quality of next generation assessments. Fordham Institute. Retrieved from:  
<https://fordhaminstitute.org/national/research/evaluating-content-and-quality-next-generation-assessments>

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>

Education First. (2023). The next generation of curriculum-connected through-year assessments. Retrieved from  
<https://www.education-first.com/innovations-in-assessment/our-impact-through-year-curriculum-connected-assessments/>

Every Student Succeeds Act (sections 1204 and 1111). Retrieved from  
<https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf>

Grissmer, D., Buddin R., Berends, M., Willingham, D., DeCoster, J., Duran, C., Hulleman, C., Murrah, W., & Evans, T. (2023). A kindergarten lottery evaluation of Core Knowledge Charter Schools: Should building general knowledge have a central role in educational and social science research and policy?. (EdWorkingPaper: 23-755). Annenberg Institute at Brown University. Retrieved from  
<https://doi.org/10.26300/nsbq-hb21>

Haertel, E. (2013). How is testing supposed to improve schooling? *Measurement: Interdisciplinary Research and Perspectives*, 11(1-2), 1-18. Retrieved from  
<https://doi.org/10.1080/15366367.2013.783752>

Hawai'i Kaiapuni Assessment of Educational Outcomes. (n.d.). Retrieved from  
<https://www.hawaiipublicschools.org/TeachingAndLearning/Testing/KAEO/Pages/home.aspx>

Healy, A. F., & Wohldmann, E. L. (2012). Specificity and transfer of learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 227-253). Elsevier Academic Press. Retrieved from  
<https://doi.org/10.1016/B978-0-12-394293-7.00006-6>

Hirsch, E. D. (2016). *Why knowledge matters: Rescuing our children from failed educational theories*. Harvard Education Press.

Hoff, D. J. (2005). N.Y. 'Portfolio Schools' get Regents reprieve. Ed Week. Retrieved from  
<https://www.edweek.org/teaching-learning/n-y-portfolio-schools-get-regents-reprieve/2005/08>

International Baccalaureate Organization. (n.d.). Retrieved from  
<https://www.ibo.org/>

Learning Policy Institute. (in prep). Design principles for instructionally relevant assessment systems.

Louisiana Department of Education. (n.d.). Louisiana's key initiatives - Innovative Assessment Program. Retrieved from [https://www.louisianabelieves.com/docs/default-source/key-initiatives/louisiana's-key-initiatives---innovative-assessment-program.pdf?sfvrsn=a6219f1f\\_24](https://www.louisianabelieves.com/docs/default-source/key-initiatives/louisiana's-key-initiatives---innovative-assessment-program.pdf?sfvrsn=a6219f1f_24)

Lyons, S., & Marion, S. F. (2016). Comparability options for states applying for the Innovative Assessment and Accountability Demonstration Authority: Comments submitted to the United States Department of Education regarding proposed ESSA regulations. Retrieved from

<https://www.nciea.org/library/comparability-options-for-states-applying-for-the-innovative-assessment-and-accountability-demonstration-authority-comments-submitted-to-the-united-states-department-of-education-regarding-proposed-e/>

National Academies of Sciences, Engineering, and Medicine. (2018). How people learn II: Learners, contexts, and cultures. The National Academies Press. Retrieved from

<https://doi.org/10.17226/24783>

National Research Council. (2000). *How people learn: Brain, mind, experience, and school*. The National Academies Press. Retrieved from

<https://doi.org/10.17226/9853>

New York Performance Standards Consortium. (n.d.). Retrieved from

<http://www.performanceassessment.org/>

NWEA. (n.d.). Louisiana IADA Brief: Informing the development of an innovative assessment program. Retrieved from

[https://www.nwea.org/resource-center/white-paper/47609/Louisiana-IADA-Brief\\_NWEA\\_whitepaper-1.pdf/](https://www.nwea.org/resource-center/white-paper/47609/Louisiana-IADA-Brief_NWEA_whitepaper-1.pdf/)

Office of Elementary and Secondary Education. (n.d.). Innovative Assessment and Accountability Demonstration Authority (IADA). Retrieved from

<https://oese.ed.gov/offices/office-of-formula-grants/school-support-and-accountability/iada/>

Office of Elementary and Secondary Education. (2018). A State's Guide to the U.S. Department of Education's Assessment Peer Review Process. Retrieved from

<https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>

Shepard, L. A. (2021). A call to action for educational researchers. *American Educator*, Fall 2021. Retrieved from

<https://www.aft.org/ae/fall2021/shepard>

Shepard, L. A., Dymoke, S., Kamenetz, A. H., Lindsay, J., Linn, R. L., & Chabay, R. (2018). Learning, motivation, and assessment: A conceptual analysis with implications for classroom practice. *Measurement: Interdisciplinary Research and Perspectives*, 16(1-2), 37-62. Retrieved from

<http://www.edmeasurement.net/MAG/Shepard-et-al-2018-EMIP-learning-motivation-assessment.pdf>

Sireci, S. G. (2020). Standardization and UNDERSTANDARDIZATION in educational assessment. *Measurement: Interdisciplinary Research and Perspectives*, 18(1), 1-4. Retrieved from

[doi:10.1111/emip.12377](https://doi.org/10.1111/emip.12377)

Slover, L. (Personal correspondence, 2023). CEO, CenterPoint Education Solutions.

VanderVeen, A. (Personal correspondence, 2023). CEO, New Meridian.

Wexler, N. (2019). *The knowledge gap: The hidden cause of America's broken education system—And how to fix it*. Avery.

Willingham, D.T. (2017). *The reading mind: A cognitive approach to understanding how the mind reads*. Jossey-Bass.

Wooldridge, C., & Weinstein, Y. (2016). What's transfer, and why is it so hard to achieve? Parts 1 & 2. The Learning Scientists. Retrieved from

<https://www.learningscientists.org/blog/2016/6/2-1>

