7

State Practices and Balanced Assessment Systems

Carla Evans, *National Center for the Improvement of Educational Assessment*

Erika Landl, *National Center for the Improvement of Educational Assessment*

CONTENTS

INTRODUCTION

In this chapter, we argue that the primary role of states in promoting balanced assessment systems should be to create and support the right structures and conditions for district and school leaders and classroom educators to effectively improve student learning. We acknowledge that states cannot design or implement balanced assessment systems on their own because they have limited control over the elements that comprise or influence local decisions. Most decisions that impact the design and implementation of local assessment systems are made at the district-, school-, and classroom-level. This is not to imply or suggest that states do not have a critical role to play in supporting more balanced assessment systems but simply acknowledges that states serve a supporting role, which represents a difference in action—not importance. The purpose of this chapter is to discuss how states can have a positive and appropriate influence on the portfolio of assessments used by educators to support ambitious teaching and equitable assessment practices.[1]

Several factors influence a state education agency's (SEA's) impact on local assessment policies and practices.[2] While some of these factors are out of a SEA's control (e.g., federal mandates), others represent a state's unique political landscape and perceived role in supporting student outcomes, both of which can influence the resources and autonomy afforded to districts and schools in making decisions about curriculum, assessment, and instruction. For example, while all states have federally required summative assessment programs,[3] some SEAs also provide districts and schools with optional interim assessments and associated resources, progress monitoring tools, and item banks. Similarly, local education agencies (LEAs) vary in the degree to which they value, trust, and use results from the state summative assessment program and/or other state-provided tools to improve local practices. In some contexts, the SEA is seen by the LEA as a crucial partner; in other contexts, the SEA fulfills only a monitoring function—and many state and local partnerships fall somewhere in between these two extremes.

Given the marginal influence that SEAs have on local assessment decisions, an SEA's effectiveness in promoting balanced assessment systems rests on its ability to create and promote structures, policies, and resources (e.g., tools, guidance) that (a) foster trust between state and local entities as well as reciprocal accountability; (b) incentivize practices that prioritize students' unique learning needs and academic outcomes; (c) signal what is important to teach and for students to learn; (d) promote fair, appropriate, inclusive, and equitable assessment practices; and (e) discourage and

[1] Ambitious teaching "centers on each student's engagement and participation; it requires paying explicit attention to who students are as they enter the classroom, including their prior learning experiences (inside and outside formal educational settings), their family- and community-based funds of knowledge, and their races, ethnicities, gender identities, social classes, and other aspects that influence their identities as learners…. Equitable assessment is embedded in and enables ambitious teaching" (Shepard, 2021). See Shepard (2021) for more details.

[2] Throughout this chapter, we distinguish the state education agency (SEA) from other state governmental bodies (e.g., state legislature, governor's office, the state board of education) that can influence or constrain decisions about the state assessment program and how it is viewed or used. In this chapter, the term "state" with no modifier refers to both the SEA and state governmental bodies.

[3] Throughout this chapter, we use the following terms interchangeably: federally required/mandated state summative testing program, state summative assessment program, state-required annual achievement tests, state tests, state testing programs, and state summative tests.

mitigate assessment practices that perpetuate systemic inequities and/or work against efforts to create rich learning environments (Chappuis et al., 2016; Conley & Darling-Hammond, 2013; National Research Council, 2010; Stiggins, 2006, 2008, 2017).

The argument laid out in this chapter rests on two overarching assumptions. The first assumption is that a state's influence on local assessment practices can and should extend beyond the state's summative assessment program. The second is that balance is best conceptualized as existing along a continuum. Each of these assumptions is discussed below.

Measures of student participation and performance on federally mandated state summative assessments are included in school accountability determinations. Consequently, factors that influence student performance on the state summative assessment are likely top of mind when school and district leaders make decisions about the materials used to drive and evaluate teaching and learning. This signaling function is beneficial to the extent that the state summative assessment clarifies the expectations underlying the state content standards, does not work against deeper learning practices, demonstrates high-quality item and test development, and demonstrates implementation practices that support appropriate test use and interpretation (e.g., attainable expectations for performance on the test have been established). However, given its influential role, the state summative assessment program can also have an outsized negative influence if not situated as one element of a broader system of assessments.

Within a school year, students participate in a broad range of assessments directed or mandated by actors at different levels of the educational system (e.g., district, school, and classroom) for different purposes (e.g., screening, instruction, evaluation). To provide information to stakeholders that will improve their decision making and positively impact teaching and learning, the assessments must work together to provide a useful and coherent profile of information about student achievement (i.e., learning strengths and needs, student performance and growth) (Marion, 2019b). Achieving this goal is no easy task—it requires the coordinated planning and engagement of multiple stakeholders and a basic understanding, at the very least, of the fundamentals of assessment design. If states do not participate in efforts to improve local assessment practices "there is a greater likelihood that assessment systems will remain incomplete or incoherent" (Gong, 2010).

A local assessment system includes all the assessments administered to students in a year in a district or school, including state-required annual achievement tests, school- and/or district-required assessments, and classroom assessments. Therefore, SEAs cannot dictate the design of local assessment systems but can provide resources and guidance that (a) clarify the intended role of the state summative assessment program and any other state-provided tools, (b) improve assessment literacy, and (c) advocate for policies, opportunities, or incentives that will foster improved local assessment practices. In some cases, SEAs will have a greater influence on local assessment practices, such as when districts and schools are identified for state support. States can also proactively address issues within their control that are likely to undermine the balance of local assessment systems, as discussed in more detail later in this chapter.

To this chapter's second assumption, balance is not dichotomous—it is a matter of degree. An assessment system cannot be unbalanced one day and balanced the next. Balance exists on a continuum, which reflects the extent to which desired

characteristics—coherence, continuity, comprehensiveness, efficiency, and utility—are represented in the set of assessments under consideration (in this volume, see Chapter 1, "Reimagining Balanced Assessment Systems: An Introduction," and Chapter 2, "The Struggle to Implement Balanced Assessment Systems: Explanations and Opportunities"). In addition, since assessments are selected or determined by a diverse set of stakeholders, balance may be differentially represented at various levels of the educational system. For example, after conducting a local assessment system audit, a district may discontinue poorly aligned assessments and/or assessments that are not perceived as providing timely, useful information. In doing so, the district will increase the overall balance of its assessment system by improving efficiency and utility, while confirming that the remaining assessments are coherent with each other and the district-defined curriculum (e.g., reflect a common approach to learning). This does not mean, however, that the district's assessment system will meet the needs of all stakeholders or reflect a clear, consistent message about where students need support when paired with classroom assessment information. In fact, the district's attempt to improve efficiency may be perceived by some as negatively impacting the comprehensiveness of the system (i.e., the range of information provided to inform decision making). A separate analysis that considers school-level assessment practices in combination with the materials and assessments enacted at the classroom level may be necessary to evaluate the degree to which balance is reflected within and across these levels and how it can be improved. Consequently, improved balance at one level of the system—whether it be state, district, or school—is not sufficient to ensure that the overall system will have a positive impact on teaching and learning.

Additionally, SEA leadership cannot articulate what a district or school leader or classroom educator needs at the level of granularity necessary to design local practices around curriculum, instruction, and assessment. However, SEAs can think about the role they believe the state summative assessment—and other state-developed and -provided tools—can or should play. The SEA can then ensure that communication and resource efforts are aligned with those beliefs to appropriately inform local assessment decisions and practices.

The efforts necessary to promote balance depend on the range of information that stakeholders need to support student learning. These needs can change over time in predictable and unpredictable ways. Therefore, in the same way assessment validation requires ongoing evidence collection to support score interpretation and use, maintaining balance should be perceived as an ongoing process of adjustment rather than an attainable end state.

The principal audience for this chapter is SEA personnel who are tasked with designing and implementing the state's vision for education through multiple means, including the state's assessment program. A secondary audience is state legislatures, state boards of education, and state chiefs, who have significant control over education policy—including policies that can afford or constrain state assessment and accountability decisions made by SEAs and LEAs. Another key audience is test vendors, who are partially responsible for operationalizing a state's vision for its assessment program or system.

This chapter begins by situating a state's role in designing and implementing balanced assessment systems within a larger sociopolitical context. Specifically, we

consider how federal accountability and peer review requirements influence state assessment decisions and exert pressure on districts and schools that can trickle down to the classroom.

Subsequently, we discuss how these contextual factors result in state actions that impact the balance of assessment systems. In particular, we focus on the outsized and often unclear way that districts, schools, and classrooms use state summative assessment results to inform decisions aimed at improving student learning. We also note the lack of systems thinking demonstrated by SEAs with respect to supporting the design and implementation of high-quality assessment practices due to concerns over local control.

Next, we compare what is under local versus state control regarding the design and implementation of balanced assessment systems. This section serves to ensure that the recommendations that follow are appropriately aligned to the decisions state education agencies, state boards of education, state legislatures, and state chiefs are tasked with and can reasonably change.

Finally, the remainder of this chapter focuses on high-leverage actions states can take to promote the design and implementation of more balanced assessment systems within and across levels of the educational system.

BACKGROUND

High-Stakes Federal Accountability

Accountability uses of state assessment results (e.g., school accountability ratings or designations) can work against state and local efforts to develop balanced assessment systems in obvious and hidden ways (Fuhrman & Elmore, 2004). Specifically, using accountability as a driver for whole-school reform can impede ambitious teaching and equitable assessment practices because the pressure to raise student test scores will take precedence over student learning. Elevating the importance of test scores can result in the proliferation of commercial interim assessments to predict performance and monitor progress (Marion et al., 2019), teaching the test (Supovitz, 2009), narrowing of the curriculum (Au, 2007), educational triage (Booher-Jennings, 2005; Diamond & Spillane, 2004), and other pernicious practices and effects (Firestone et al., 2000). Consequently, state efforts to positively impact local assessment practices will not be successful without thoughtful reform in how federal accountability is enacted.

However, accountability is not always negative. For example, adding state summative tests in social studies can help balance the incentive to focus solely on English language arts, mathematics, and science to the exclusion of other core subject areas. Additionally, accountability regulations that require school-level academic achievement to be disaggregated by student group can highlight the achievement gaps of marginalized communities and under-represented groups like students with disabilities and English learners. For example, as advocates from these student groups would contend: "we cannot fix what we cannot measure" and their organizations "rely on the consistent, accurate, and reliable data provided by annual statewide assessments to advocate for better lives and outcomes for our children" (The Leadership Conference on Civil and Human Rights, 2015).

Federal Requirements

Federal regulations directing the design, administration, reporting, and peer review[4] of state assessments create conditions that influence the role the state summative assessment can and should play in districts and schools. For example, federal requirements dictate annual state testing for students in certain grades and subject areas. These state tests must meet certain requirements related to validity, reliability, and fairness, and must also produce individual student reports that allow stakeholders to understand and address students' specific academic needs. These requirements, which are evaluated as part of the federal peer review process, are designed to ensure that the state summative assessment provides high-quality information to inform public reporting and support school accountability.

Due to these requirements, states have specific constraints that affect how they design, deliver, score, and report their summative assessments. These constraints and their impacts on the assessments may send unintended signals to local educators about how learning is best evaluated. For example, federal peer review requires states to submit evidence that assessments have been designed to support student proficiency on the breadth and depth of grade-level academic content standards, comparable across classrooms in the state. Accommodating this requirement necessitates a content sampling design that meaningfully represents the grade-level standards and supports the development of items that can be evaluated within the context of a large-scale standardized assessment (e.g., selected or short-answer responses that can be scored accurately and consistently). These prioritizations, reflected in the design of the summative assessment, can negatively influence how and what teachers teach if not accompanied by clear communication about the role and purpose of the state summative assessment and the rationale underlying its design.

Despite the challenges that the peer review process can create for state summative assessment design, the process can also have a peripheral, positive impact on efforts to support balanced assessment systems. Currently, peer review is the only process that exists for evaluating the technical quality of state-designed assessments in a comprehensive, standardized manner. Because it serves as an independent standard for quality, peer review is a useful criterion when working with state governmental bodies that may want to implement assessment policies that could undermine technical quality and inclusivity because it can be used as an argument against such policies. Furthermore, because peer review results are used to label schools, identify them for support, and inform other state decisions (promotion, grades, teacher evaluation), ensuring a state's summative assessment demonstrates technical quality is a necessity.

Outsized Role of State Assessments in Shaping Local Curriculum, Instruction, and Assessment

State assessments take an outsized role in local contexts when educators begin reshaping local curriculum, instruction, and assessment to mimic the format or struc-

[4] It is beyond the scope of this chapter to discuss the state assessment peer review process in detail. We refer readers to the U.S. Department of Education's guide to the peer-review process here for additional context: https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf.

ture of the state test—especially for state tests that are mainly selected responses. The state test is one instantiation of state content standards—and a limited one, given the design features necessary to fulfill federal accountability requirements (e.g., the standardization necessary to support comparability). Additionally, political, logistical, and practical constraints often limit what can be assessed—and at what depth—in state assessments. Therefore, while the content of the state test should be mirrored in the local curriculum and instructional program because both are built from the same set of content standards, the format and structure of classroom assessments can—and likely should—be more varied and distinct from the state tests. The inclusion of performance tasks on state tests is one way that states can use the often outsized role of the state test to signal the importance of complex demonstrations and applications of learning at the local level.

STATE ACTIONS CONTRIBUTING TO IMBALANCED ASSESSMENT SYSTEMS

States contribute to imbalanced assessment systems by perpetuating or inflating structures and conditions that work against efforts to administer local assessments that complement curriculum and instruction. Imbalance can occur for many reasons, including layering on additional high-stakes accountability decisions based on state assessments, unclear communications about how state assessments will be used to inform decisions or actions, and failure to envision state assessments within the broader system of assessments.

Layering on Additional High-Stakes Accountability Decisions Based on State Assessments

SEAs must adhere to federal school accountability guidelines. However, many state governing bodies go beyond federal rules and regulations, requiring the use of state summative assessment results for additional high-stakes decisions. For example, some state legislatures require student test scores to be part of teacher evaluations, even though the state test is not designed to support inferences about teacher effectiveness (American Educational Research Association, 2015). In other states, state summative assessment results have been used to determine high school graduation, third-grade promotion based on literacy performance, and other gateway decisions. Using student test results in these ways is not required under federal law and can lead to behaviors, practices, and conditions that work against efforts to support balanced assessment systems. This is not to say that all additional state accountability leads to imbalance. It depends on who is being held accountable, for what, and the evidence or theory that supports such actions.

Unclear Communications About How State Assessments Will Be Used to Inform Decisions or Actions

Each year SEAs administer tests to all students in federally required grades and content areas. Although SEAs spend exorbitant amounts of time and money ensuring

that these assessments meet federal peer review requirements, few SEAs provide a theory of action that describes, in detail, how the state assessment program is intended to drive progress or inform decisions that positively impact school quality and student outcomes. This lack, coupled with a dearth of clear communication about the primary purpose of state testing, perpetuates misconceptions about how assessment results can and should be used (e.g., by teachers to make instructional decisions). These misconceptions, in turn, fuel concerns about over-testing and the value of the state assessment when those other desired uses, such as instructional usefulness, are not supported. Dissatisfaction with state assessments and opt-out movements are the visible signs of these fractures.

Similarly, states often create policies or initiate assessment reforms that influence the design, use, or impact of the state assessment program absent a clear theory of action that defines how the changes will lead to improved teaching and learning. This lack of a clear theory of action is reflected in how some states are considering or piloting through-year assessment designs. Dadey and Gong (2017, 2021) define a through-year assessment program as having assessments that are (1) administered in multiple distinct sessions during a school year, and (2) intended to support the production and use of a summative determination of student proficiency and one additional aim. The additional aim is often instructional utility. In essence, these through-year assessment reforms are trying to make state assessments serve multiple roles—the typical monitoring and accountability role, which is federally required, and an instructional support role. What is often left underspecified, however, is how the information supplied by the through-year state assessments (e.g., raw score, achievement level, scaled score) will foster high-quality instructional actions and practices at the local level (Dadey et al., 2023). For example, what specific action(s) does the state expect classroom educators to take with the assessment results? How does the grain size and frequency of information provided serve to support that use? What assumptions must hold for it to do so effectively?

Failure to Envision State Assessments Within the Broader System of Assessments

When SEAs focus their efforts and communications solely on the state summative assessment program, they give up a powerful opportunity to help stakeholders understand and appropriately situate the state test within the broader system of assessments used to collect information about student performance over a year. Choices made at one layer of the assessment system can have a trickle-down or filter-up effect that can drive imbalance by constraining or inappropriately influencing decisions and actions. States can help facilitate systems thinking by (a) modeling this practice when communicating about the intended use of state assessments in relation to locally administered assessments and (b) providing assessment design and evaluation tools, resources, and supports that promote systems thinking. Each of these topics is discussed in more detail in the section titled "State Actions to Support Balanced Assessment Systems."

Additionally, state laws or policies related to assessment can work against teaching and learning and signal different instructional priorities from that of the content standards. For example, laws or policies that focus on or necessitate keeping state tests short

and cheap (e.g., state-defined constraints on testing times and the federal requirement to test every student every year) could lead to decisions like the elimination of writing prompts or more complex item types. However, these types of items elicit students' knowledge and skills related to the depth of the content standards. Such policies or lack of funding could result in state tests that do not appropriately signal instructional priorities around deeper learning and work against the models of learning that the SEA is trying to promote. This disconnect also exemplifies a lack of systems thinking. The state test should reflect the content standards and how they are intended to be taught. For better or for worse, what gets tested gets taught—at least to some degree (Faxon-Mills et al., 2013).

BALANCED ASSESSMENT SYSTEMS: WHO CONTROLS WHAT?

Before focusing on SEA actions that support balanced assessment systems, it is important to differentiate what is directly in the control of the state versus what is directly in the control of LEAs. *Education is a federal interest, a state responsibility, and a local function*. So, who has control of what aspects of the educational system when decisions need to be made about the factors (e.g., education policies, resources, and actions) known to indirectly or directly influence balanced assessment systems? It is important to clarify these roles upfront so that the recommendations in this chapter for SEA actions align with what is under state control.

Local Control

As shown in Table 7-1, LEAs control decisions about curriculum, instruction, and local assessments, as well as defining local assessment policies and practices (see Chapter 6 of this volume, "District and School Practices and Assessments to Support a Learning-Centered Vision"). For this reason, it has been stated that the primary locus of control for the design and implementation of balanced assessment systems lies with

TABLE 7-1 Areas Under State Versus Local Control Related to Balanced Assessment Systems

State Controlled	Locally Controlled
• State content and performance standards • Federally required state summative assessments • School accountability systems (e.g., school ratings or rankings) • Teacher standards, licensure, and recertification • Educator preparation program approvals (initial and ongoing) • Additional state-required assessments (e.g., social studies state assessments, dyslexia screeners, universal screeners) or other state-provided assessments (e.g., optional interims) • State supports, guidance, tools, and/or resources offered to local education agencies around curriculum, instruction, and assessment	• Local curriculum, instruction, and assessment decisions, including school- or district-required assessments and classroom assessments • Selection and implementation of professional learning opportunities for teachers • Local assessment policies and practices (e.g., grading policies and requirements regarding curriculum pacing and scripting)

LEAs (Marion, 2018; Marion et al., 2019; Shepard et al., 2018). This is not to say that decisions at the state level do not affect local actions—the accountability function and design of state summative assessments can incentivize or undermine local efforts to design and implement balanced assessment systems. However, decisions about curriculum, teaching and learning priorities, classroom assessment strategies, and local measures of student progress fall within the purview of local school boards, district and school leaders, and teachers.

Decisions made at the LEA level can have a significant impact on teachers' classroom activity systems (see Chapter 4 of this volume, "Classroom Activity Systems to Support Ambitious Teaching and Assessment"). Over-testing in schools is often due to local testing requirements, although the blame is often laid on the state (Marion et al., 2019).[5] For example, LEAs may require teachers to administer universal screeners, literacy assessments, benchmark assessments, and/or commercial interim assessments beyond those required by the state to track student progress and gauge proficiency. These LEA-selected and -required assessments, in addition to school and teacher assessment preferences, can cause over-testing at the local level, as well as an overreliance on standardized measures of student performance to inform educators' instructional practice. While the prevalence of over-testing is true for all students, it is even more serious for English learners, who often need to take both state- and locally mandated English proficiency tests in addition to all other assessments.

LEAs also control the selection and implementation of professional learning opportunities for teachers. Designing and implementing balanced assessment systems requires educators to have strong pedagogical content knowledge and assessment literacy. Therefore, LEAs need to provide related opportunities for professional learning and capacity-building resources to encourage improved implementation of high-quality local assessment systems (in this volume, see Chapter 5, "Assessment Literacy and Professional Learning," and Chapter 6, "District and School Practices and Assessments to Support a Learning-Centered Vision"). These professional learning opportunities may come from a variety of sources, including regional laboratories and support structures, external professional development providers, and state-provided training.

State Control

As shown in Table 7-1, states have control over a broad array of factors that can positively or negatively shape local systems of assessment, the quality of the information they yield, and how the information is used. The decisions a state makes regarding many of these factors—particularly additional state-required assessments and state supports, guidance, tools, and/or resources—will depend on its capacity, as well as its vision for teaching, learning, and assessment.

State content and performance standards, federally required state summative assessments, and school accountability systems are constrained by federal statutory regulations and guidance, but the state still has considerable influence over the standards, assessment, and accountability landscape. Specifically, the state—via content

[5] One caveat to this statement is the movement of some states toward mandating universal screeners, literacy assessments, or other types of assessments multiple times a year. We express caution about this movement in the section of this chapter titled "State Control."

standards—specifies what is taught, at which grade levels, and at what level of cognitive rigor for all core subjects. Similarly, the annual state test serves to evaluate how successful schools have been in supporting student attainment of identified standards. While federally mandated state summative assessments must meet certain technical requirements, the state has latitude in test design, reporting and administration features (e.g., item types, cognitive rigor, test length), performance standards, and whether to test grades and subjects beyond federal requirements. The state evaluates schools in accordance with the rules of its accountability system. State tests and the associated accountability rules can have an outsized role in shaping local teaching practices and curricula, both positively and negatively.

Because states control teacher standards, licensure, and recertification, SEAs could require educators to meet state-defined assessment literacy requirements to receive a teaching license. Similarly, concerning recertification, the state could offer continuing education credits, micro-credentialing, or badging options related to the demonstration of assessment literacy.

The SEA is also responsible for approving educator preparation programs, both initially and on an ongoing basis—a lever it can use to incentivize or mandate coursework and clinical experiences consistent with the state's theory of action around balanced assessment systems. We discuss what this type of professional learning might entail and how the SEA could support such efforts later in this chapter.

State boards of education and legislatures decide what, if any, state-defined assessments must be implemented beyond those required by federal law. Examples include K–2 literacy screeners, universal screeners, social studies state assessments, high school end-of-course exams, and additional science testing (i.e., beyond once per grade span). Some SEAs also supply districts and schools with optional interim assessments—either created by the state or purchased through a commercial vendor—that are aligned to the state's content standards and performance level descriptors. States should be wary of contributing to the possible incoherence and inefficiency of local assessment programs and the over-testing of students when layering additional assessments on top of what is already federally required without a clear theory of action and rationale.

States also determine the types of support, guidance, tools, and/or resources offered to LEAs regarding the implementation of high-quality instructional materials and local assessment practices. Some states provide more support and resources than others due to differences in capacity and vision. In any case, the SEA can play an important role as a convener and connector of LEAs to share best practices. We discuss this issue at length in the section titled "State Action 5: Provide Tools, Resources, and Supports to LEAs."

STATE ACTIONS TO SUPPORT BALANCED ASSESSMENT SYSTEMS

This section highlights six high-leverage actions SEAs can take in supporting local efforts to design and implement balanced assessment systems (see Figure 7-1). Each action builds from one or more of the state-controlled factors and how the state wants to support locally controlled factors, if at all. Although we acknowledge that it is not yet fully known how these actions interact, we suspect that SEAs must attend to all six actions in some manner to adequately support balanced assessment systems.

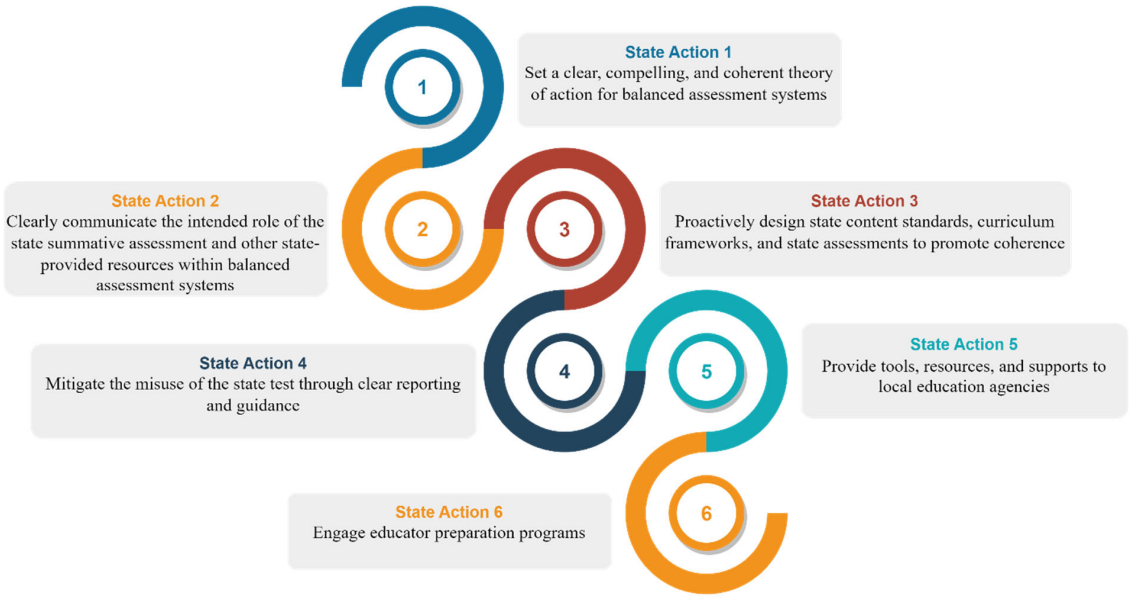


FIGURE 7-1 Six high-leverage state actions to support balanced assessment systems.

State Action 1: Set a Clear, Compelling, and Coherent Theory of Action for Balanced Assessment Systems

The first action—setting a clear, compelling, and coherent theory of action for balanced assessment systems—is the glue that holds all the high-leverage actions together. A state's theory of action for balanced assessment systems describes how the SEA understands assessment's role in supporting teaching and learning, as well as the conditions that must be in place for it to fulfill that role. This theory of action should include assessments required or offered by the state, in addition to those required locally or implemented by educators in the classroom.

Supporting the implementation of balanced assessment systems represents one of several ways a state can champion its educational vision for students. Other avenues include the design of school accountability systems, the development of state policies that influence how and when students learn, and defining course and graduation requirements. To ensure these different approaches represent a thoughtful, coherent strategy rather than a variety of disparate initiatives, a state's vision should articulate (a) the educational outcomes required to realize the vision (e.g., measures of academic and non-academic performance, participation in extracurricular activities, performance on college and career-ready assessments, acceptance into college or a vocational program upon graduation) and (b) the way those outcomes are likely to be met. Specifically, the theory of action should define the experiences and learning opportunities perceived as necessary for students to achieve these outcomes, as well as the necessary structures, interactions, and information for schools and educators to effectively incentivize and support those opportunities.

A SEA's theory of action for balanced assessment systems should describe the type of assessment information needed by different stakeholders and how the information

gleaned should be prioritized and utilized to positively impact teaching and learning consistent with the state's vision. Specifically, what information should the state assessment program provide to help achieve the state's educational goals? What information should be generated by other levels of the system (district, school, and classrooms)? How can the SEA help ensure that the state's assessments work together—and not at cross purposes—with district, school, and classroom assessments, supporting rich learning environments?

While SEAs may differ in how they support or enforce key aspects of their theory of action, all SEAs should clearly and consistently describe the role of the state summative assessment. This description should include both how results should be used and how decisions about assessment design are intended to influence what happens in schools. Since the primary purpose of the state summative assessment is to monitor and evaluate school quality, there is no direct link between the information afforded and how to improve teaching and learning practices. The information gleaned from the state summative assessment is too distal from instruction, not at the right grain size, and not timely enough to shape the daily interactions of teachers, students, and the content (Evans & Marion, in press). Furthermore, as discussed in Chapter 4 of this volume, "Classroom Activity Systems to Support Ambitious Teaching and Assessment," these assessments do not account for the classroom learning environment or consider students' cultural and social backgrounds and funds of knowledge, so they cannot provide information to support ambitious teaching. Future state assessments could be designed to account for learner characteristics, but more personalized and flexible approaches to state standardized assessment do not currently exist (Buzick et al., 2023).

However, the state assessment does provide useful aggregate data for school and district leaders to consider when making programmatic decisions such as how to allocate resources and support, the need for curriculum and staffing modifications, and the effectiveness of new initiatives or programs. In addition, state test data can and should have a positive—albeit indirect—impact on teaching and learning, consistent with the intended role of the assessment in the SEA's theory of action. Table 7-2 reflects two roles and associated theories of action for how a state's summative science assessment design and/or associated resources may indirectly influence teaching and learning.

As shown in these examples, the theory of action reflects an assumption that state assessment design and resources will promote actions or practices that ultimately lead to improved teaching and learning. As with any theory of action, steps must be taken to ensure that these assumptions hold (e.g., that the tools and resources are useful, that educators have time to work together, that sample materials and tasks are high quality) and the desired impact is realized.

Any theory of action that over-emphasizes the role of the state summative assessment program in supporting teaching and learning is bound to cause imbalance. The SEA can support more balanced assessment systems by filling in the missing links from system components such as state content standards and performance expectations to mechanisms that lead to systemic improvement and change. A key aspect of a state's theory of action for supporting balanced assessment systems is therefore clarifying how the state assessment program should inform or work with classroom curriculum, instruction, assessment, and other system components to promote student learning (see also State Action 3 later in this chapter).

TABLE 7-2 Two Abbreviated Example Theories of Action for State Science Assessment Design and/or Associated Resources

What Is the Role of State Summative Assessment?	How Will the State Assessment Positively Impact Teaching and Learning (Abbreviated Theory of Action)?	What Are the Implications of These Decisions for State Assessment Design and Other Necessary Resources?
Example 1		
To signal the type of authentic, complex tasks students should be able to engage with to demonstrate science learning as envisioned in the Next Generation Science Standards (NGSS).	It will influence how educators engage with students and each other to teach and evaluate the attainment of science expectations within and across years.	Assessment must include one or more high-quality authentic performance tasks, which may impact the time it will take students to complete the assessment and the cost of test design and scoring.
To provide data that help schools and educators evaluate how well existing curriculum and instruction prepared students to generalize their learning to novel tasks.	It will help design curriculum and instruction materials that focus on how to identify and solve authentic problems rather than only the attainment of discrete knowledge and skills.	The state must provide resources (e.g., sample tasks, scoring rubrics) and training that will help educators prepare students for success.
Example 2		
To clarify how the expectations reflected in the NGSS are distributed and evaluated within and across grades given the state's vision for science education.	It will support schools and districts in establishing a strategy for addressing NGSS expectations within and across grades.	The state must provide resources (e.g., released items) and training that clarifies how the expectations underlying the standards should be addressed and evaluated within and across grades.
To provide data that allows schools and educators to evaluate how well existing curriculum and instruction prepared students to meet expectations at the end of a particular grade span.	It will allow districts and schools to collaborate in the development of curriculum, instruction, and assessment resources based on a shared understanding and trajectory for science attainment.	There is a need to create summative assessment frameworks that complement these resources and reflect priorities for monitoring and evaluating performance (e.g., reporting categories).

If a SEA provides other assessment-related tools, resources, and support—see also State Action 5 later in this chapter—its theory of action should explain how they are intended to support the state's educational goals. For example, if the SEA provides free assessment literacy resources to all classroom educators and school and district leaders, then the theory of action should explain how providing those free resources is logically connected to improving classroom instruction and assessment practices.

Figure 7-2 provides one hypothetical depiction of a state's theory of action for balanced assessment systems. The state-controlled assessment components include state-required accountability assessments as well as other state-provided—yet optional—interim and classroom assessment tools, resources, and support (see the yellow box in Figure 7-2). The depiction is meant to communicate how state-required assessments are intended for the limited purpose of program quality monitoring and evaluation. Yet because this hypothetical SEA wants to promote balanced assessment systems, providing interim and classroom assessment tools, resources, and supports can promote and support the quality of local assessments in the teaching and learning feedback cycle. The locally controlled assessment components are in the green box in Figure 7-2.

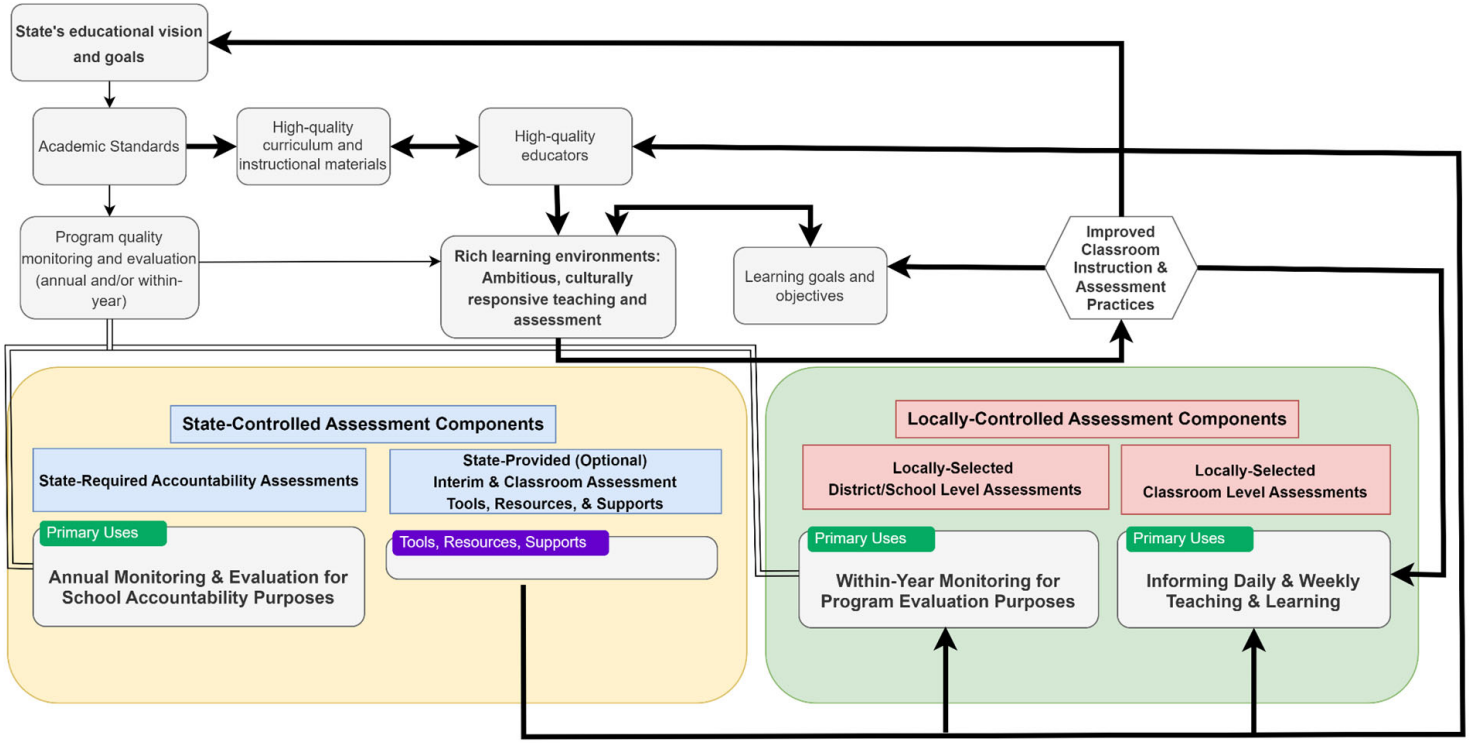


FIGURE 7-2 One hypothetical depiction of a state's theory of action for balanced assessment systems.

It is beyond the scope of this chapter to go into detail about the theory of action, but what should be clear from Figure 7-2 is that there is a complex array of interactions expected and desired among the (a) state-provided tools, resources, and support; (b) locally controlled assessment components; (c) rich learning environments; and (d) improved classroom instruction and assessment practices. The ultimate objective of the theory of action is to support the state's educational vision and goals for its graduates while recognizing that other offices or departments at the SEA—as well as local educational systems—are also working toward the same vision and goal.

We did not include the hypothetical example to suggest that every SEA needs to provide a similar level of support for locally controlled assessment components. Every SEA has unique concerns, needs, and capacities that drive which goals and problems are prioritized and consequently what solutions are perceived as most likely to support intended outcomes. These unique needs are where identifying general constraints and requirements related to state laws and court rulings, federal laws and regulations, state historical considerations, student and school demographics, fiscal constraints, and/or capacity constraints and limitations is critical, as these constraints directly affect local solutions.

State Action 2: Clearly Communicate the Intended Role of the State Summative Assessment and Other State-Provided Resources Within Balanced Assessment Systems

A SEA can have a clear, compelling, and coherent theory of action for balanced assessment systems and nonetheless fail to communicate it to stakeholders. Communication is the bridge from ideas to implementation—it conveys the rationale for and motivation behind the SEA's decisions regarding the state summative assessment program and the provision of additional assessment resources (see also State Action 5 later in this chapter). Communication builds awareness and buy-in and educates stakeholders about the meaning and value of balanced assessment systems.

The announced purpose of the state's summative assessment program is one of the most important messages any SEA can communicate to its stakeholders (e.g., parents, students, classroom educators, school and district leaders, state policymakers, and the public). We have noticed confusion and misconceptions about state summative assessment programs that are due to a lack of accurate messaging. For example, there are abundant mixed messages about how teachers should use state summative assessment results to inform instruction. State leaders, among others, tend to promote the usefulness of their state's assessments. However, in doing so, they often, if unwittingly, overstate the instructional value teachers can derive from the quantitative results, given the accountability demands and associated design limitations of the assessments (Coburn & Turner, 2011; Datnow & Hubbard, 2015; Evans & Marion, in press; Faxon-Mills et al., 2013).

State leaders must clearly communicate the limitations of the state summative assessment program, particularly for informing classroom instruction. State test information is not useful for informing daily or weekly classroom decision making, including what teaching and learning experiences must be adapted to better meet students' specific learning needs (Shepard, 2021; Shepard et al., 2018). The timing of

state assessment information, unclear relationship to the enacted curriculum, and grain size of received information (e.g., scale score and achievement level) do not support direct instructional relevance for modifying or adapting teaching and learning practices in real-time (Evans, 2022; Evans & Marion, in press; Faxon-Mills et al., 2013; Marion, 2019b; Shepard, 2021; Shepard et al., 2018).

State, district, and school leaders must clearly and consistently specify the intended purpose and use of state assessments as per federal law—namely, monitoring and evaluating school quality for accountability purposes in a way that is comparable across schools and districts in the state. State test results allow SEAs and LEAs to monitor achievement trends and gaps; examine the efficacy of interventions, programs, and curriculum materials; direct resource allocation; and identify new and promising practices. State test results are useful for these purposes because state tests provide a reliable information source that is comparable over years and is available for every student in the tested grades and subjects.

Connecticut, for example, used its 2022 state test data and National Assessment of Educational Progress scores to identify a state-wide weakness in middle school math performance due to the education interruptions during the COVID-19 pandemic. As a result, the Connecticut State Department of Education designed a competitive grant program for the 2023–2024 school year that creates an intensive mathematics tutoring program for students in Grades 6–9, including funding and a vetted list of approved tutoring providers (The Office of Governor Ned Lamont, 2023). The SEA will then use an education research collaboration it established with institutions of higher education across Connecticut (State of Connecticut, 2023) to monitor the effectiveness of this tutoring program using state test data from participating schools.

Connecticut's clear communication about the intended purpose and use of the state assessments should also transfer to any state-provided assessment or assessment-related resource curated by the state without additional cost to LEAs (see also State Action 5 later in the chapter). The key point here is that the SEA must articulate to stakeholders why they have provided these tools and resources, their intended uses and users, and how they can support or work against more balanced assessment systems.

As with state assessment information, SEAs bear the responsibility of explaining the intended use of any provided or required assessments. For example, if the intended use is to support program evaluation conducted by school and district leaders, then the state must show the chain of reasoning, assumptions, mechanisms, and professional learning that connects the information gleaned from the assessment to that use. Doing so increases the likelihood that the assessment information will be used to make decisions that provide for better student learning, as well as preventing misuse, incoherence, and over-testing.

State Action 3: Proactively Design State Content Standards, Curriculum Frameworks, and State Assessments to Promote Coherence

A SEA does not promote balanced assessment systems as an end in and of itself; rather, many SEAs want to support and incentivize a robust vision of teaching and learning, academic achievement, and inclusive educational practices for all students.

A SEA can use the levers within its control to further this vision, including supporting the design and implementation of rigorous content standards; curriculum frameworks; state test designs with appropriate accessibility features; and other tools, resources, and support that support high-quality local curriculum, instruction, and assessment practices.

For example, one input from the SEA into balanced assessment systems is the state summative assessment program. States have considerable latitude in terms of how they design their state assessment program, as long as the tests meet federal peer review requirements. One aspect of meeting federal peer review requirements is to demonstrate that the assessment adequately represents the depth and breadth of the state content standards. Consequently, the design of the state content standards and associated curriculum frameworks can help promote coherence among curriculum, instruction, and assessment at the classroom and state levels. *Knowing What Students Know* (National Research Council, 2001) indicates that coherence is demonstrated when assessments within a system are linked through a clear conceptual base and specification of learning targets. Below, we argue that a SEA can promote coherence by proactively designing state content standards, associated curriculum frameworks, and state tests in a way that signals valued instructional priorities.

Design of State Content Standards and Associated Curriculum Frameworks

State content standards can shape teaching, learning, and assessment in classrooms because they frame what is important to know and be able to do in a specific content area at the end of each grade. State content standards underlie decisions about the design of the state assessment program and are the basis against which decisions about the quality and appropriateness of local assessments are made (e.g., alignment to standards). State content standards serve as a through line that extends from the state to the classroom and consequently play a large role in ensuring coherence (see Chapter 1 of this volume, "Reimagining Balanced Assessment Systems: An Introduction"). Currently, most state content standards are long lists of discrete knowledge, skills, and understandings that are isolated from other content within and across domains, silent on intended generalization and use, and not developmental. One way that states can promote more coherence and potentially advance impactful use of state assessment data is to write richer learning expectations for students—connecting competencies with other content within and across domains in desired ways, explicitly stating intended generalization and use, and displaying developmental structure and sequences.

States must use research on the way students learn and demonstrate more sophisticated knowledge and expertise within a domain to design state content standards and supplementary documentation and guidance, like curriculum frameworks to support ambitious teaching and equitable assessment practices. For example, the Next Generation Science Standards intentionally tried to reshape curriculum, instruction, and assessment in K–12 science classrooms to better reflect research on the ways students learn science (National Research Council, 2012). Similarly, the Common Core State Standards in math signal to the field that mathematics education is more than just procedural skill and fluency and "build on the best of existing standards and reflect

the skills and knowledge students will need to succeed in college, career, and life" (Common Core State Standards Initiative, 2021).

Some states, including California, have helped teachers understand how to implement content standards by creating curriculum frameworks (California Department of Education, n.d.). These frameworks help teachers faced with lists of discrete standards understand how to cluster and group the standards for instructional purposes, such that teachers can focus on the big ideas that are important at each grade level and recognize the underlying learning progressions tying the standards together. Understanding the underlying progressions supports accelerated learning and other differentiated approaches to instruction because it gives teachers both a heuristic for interpreting evidence of student learning and knowledge about what instructional moves are most likely to help students progress toward proficiency.

State Test Design

State tests serve as examples of inclusion practices both from a participation perspective (every student in federally required grades and subjects is counted in participation rates) and from an accessibility perspective (every student can show what they know and can do). Additionally, SEAs operationalize and demonstrate their values about how students learn a discipline and what instruction should be prioritized through the design or adoption of state tests (National Research Council, 2001, 2003). Ideally, states should design or adopt—in the case of assessment consortia—their state assessment program with a clear understanding of (a) how state tests will promote the instructional priorities that the state values and wants to see implemented; and (b) an understanding of how the state summative assessment should complement information collected through local assessment systems. Furthermore, states should strategically engage a diverse array of stakeholders in the assessment design and specification process, including those who represent the cultural, ethnic, racial, and special populations present in the state. Involving stakeholders from the beginning of the state assessment design process helps ensure the cultural validity of assessment results (Shultz & Englert, 2021) and models the type of stakeholder engagement and inclusivity desired at the local level.

States that are part of assessment consortia such as Smarter Balanced will have additional layers of complexity to consider. For example, assessment consortia, by design, somewhat constrain individual state decisions because a set of distinct state testing programs is replaced by one collective consortium testing program. States may have some leeway to adjust the test blueprint and reporting structure, but there are limits to what they can personalize when they are part of consortia.

Per State Action 1, presented earlier in this chapter, the state assessment program should be designed with a clear understanding of how it should support better educational decision making—ultimately supporting student learning—and the influence it is likely to have on local curriculum, instruction, and assessment practices. For example, part of the design work for a state assessment is considering tradeoffs associated with different test design features including item types, content priorities, adaptive testing models, and length of the test. Including longer constructed-response items and performance-based tasks on state assessments can signal the importance of cognitively

rigorous teaching, learning, and assessment. However, adding more complex item types usually results in a longer test and more expensive scoring. To promote coherence, the design of the state tests should send a clear and consistent message about what is important for teachers to teach and students to learn, and at what level of cognitive rigor.

Figure 7-3 shows a released item from the Spring 2022 Grade 10 Mathematics state test in Massachusetts (Massachusetts Department of Elementary and Secondary Education, 2023). Figure 7-4 shows an Algebra I performance task from the Mathematics Assessment Project (Mathematics Assessment Resource Service, 2011). Imagine a state test that contains only selected-response items like those found in Figure 7-3 versus a mixture of item types, including performance tasks like the one found in Figure 7-4. What view of human learning and development (see Chapter 3 of this volume, "Human Learning and Development: Theoretical Perspectives to Inform Assessment Systems") and associated classroom activity systems (see Chapter 4 of this volume, "Classroom Activity Systems to Support Ambitious Teaching and Assessment") do these different test designs support or constrain? Our point here is simple: SEAs can harness the signaling function of the state test to either promote or work against its vision of teaching and learning.

Whatever decisions are made by the SEA about the design of the state test should be clearly communicated to LEAs through blueprints, guidance documents, item samples, released annotated items, and practice tests, among other resources. To ensure that the state assessment program and associated resources reflect the state's theory of action, SEAs must clearly articulate their requirements when they release a request for proposals to vendors, including the specific claims, interpretations, and uses the state summative assessments and any additional state-provided assessments must be designed to support. Test vendors are responsive, not vision-casting entities. A vendor's job is to design assessments that reflect the state's goals, vision, and theory of action—so these must first be defined by the SEA. This is especially relevant if a SEA is looking for its state assessment program to include novel elements.

In addition, state tests should be designed, and achievement levels set, based on realistic and attainable performance expectations. Realistic and attainable expectations are essential if state tests are to have a positive influence on educational decision making and student learning. Unattainable expectations can undermine motivation and encourage inappropriate test preparation and use.

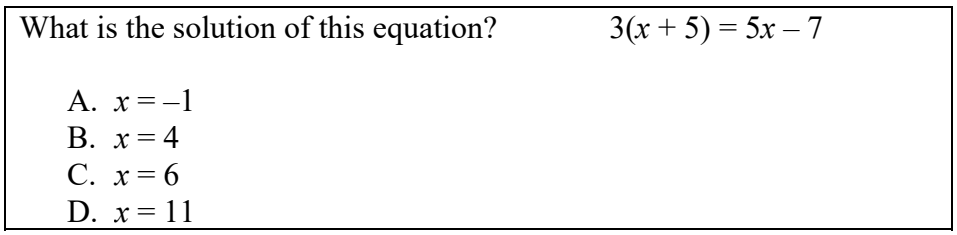


FIGURE 7-3 Selected-response item from MCAS 2022 grade 10 mathematics released items.
SOURCE: Massachusetts Department of Elementary and Secondary Education (2023).

On the grid are eight points from two different functions.

A certain linear function passes through exactly four of the points shown.

A certain quadratic function passes through the remaining four points.

For the **linear** function:

1.Write the coordinate pairs of its four points.

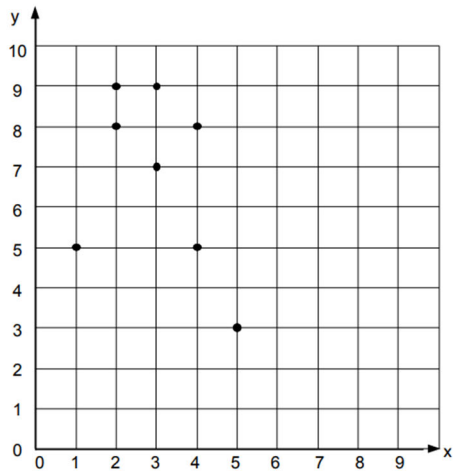
_____

_____

_____

_____

Draw the line on the grid.



2.Write an equation for the function.

Show your work.

_____

For the **quadratic** function:

3. Write the coordinate pairs of its four points.

_____ _____ _____ _____

Draw the graph of the function on the grid.

4. Write an equation that fits the quadratic function. _____
Show your work.

FIGURE 7-4 Algebra I performance task from the Mathematics Assessment Resource Service Task Bank. SOURCE: Mathematics Assessment Resource Service (2011).

State Action 4: Mitigate Misuse of the State Test Through Clear Reporting and Guidance

State test misuse can unbalance assessment systems at the local level. For example, state test results are sometimes inappropriately used to screen students out of algebra or place students into remedial coursework or non-flexible groups or tracks. Similarly, while state test results provide highly reliable information that could be used to monitor school improvement efforts over time, some districts and schools administer additional assessments for the same purpose because local users are not aware of how to use state assessment results in this way. States can mitigate state test misuse—or lack of use—through clear reporting guidance directed to users of the assessment system. These mitigation strategies promote the utility and efficiency of the entire system, which are key characteristics that support balance.

State Test Reporting Features and Guidance for Interpretation and Use

Reporting is often an afterthought or post-hoc activity following assessment design and data collection. And yet, how state assessment information is communicated to stakeholders—parents, students, classroom teachers, school and district leaders, and the public—and ultimately interpreted and used can significantly impact how stakeholders engage with, perceive, or value information from the state test or other state system components. Balanced assessment systems should provide their many diverse stakeholders with the information they need to make timely, accurate educational decisions that ultimately support student learning. Score reports are the vehicle for communicating the test's results to stakeholders and therefore must be crafted with care.

To support the utility of state information, score reports and associated resources should be designed with specific users in mind. Currently, many state-produced reports lack clear user guidance (e.g., suggested actions for school and district leaders, classroom educators, or parents), which could leave system users to interpret and use assessment results inappropriately. However, the adequate and appropriate interpretation of test scores—let alone moving from interpretation to actionable next steps—requires a high level of assessment literacy, time, and effort. Ensuring adequate and appropriate interpretation of scores might be better met, and time and effort better spent, if states provided a selection of high-impact reports that presented student and aggregate test results in multiple ways with a few high-leverage actions that different system users could take based on the results. For school and district leaders, these high-leverage actions might include gathering more contextual information on program implementation and teacher curriculum supplementation. The state likely has historical data on students, schools, and districts. This information could be used to create a reporting system that provides district- and school-level reports summarizing trends in overall student performance and reporting category for each grade and content area. This type of reporting system could help save local leaders' time, as they might otherwise try to create these reports on their own. Additionally, the reporting system might propose questions for the LEA to investigate based on student performance and trends over time.

Additionally, as with test design, states demonstrate their values through choices reflected in the state reporting system. Score reports and other assessment-related

guidance often use reporting structures and deficit-based labels—for schools and students—that can lead to interpretations that promote systemic inequities. For example, labels used to categorize student performance as 'failing' or 'well below proficient' can influence perceptions of ability and consequently teaching practices in ways that contradict rich and inclusive learning environments.

State Action 5: Provide Tools, Resources, and Support to LEAs

As stated throughout this chapter, states have little, if any, control over the composition and implementation of local assessment systems. However, SEAs can influence and promote high-quality assessment policies and practices at the local level by directly creating or curating tools, resources, and support for assessment-related endeavors. SEA personnel can further their work in these areas by attending professional conferences and meetings, as well as engaging in professional networks where they can learn from other SEAs, researchers, practitioners, and organizations. We discuss tools, resources, and support created by SEAs as one of the final state actions because they must flow from the state's vision, associated communication strategies, and other proactive and mitigation activities related to the state assessment program.

In this section, we discuss the assistance states should provide LEAs to some degree to inform the design and implementation of balanced assessment systems at the local level. However, we recognize that SEA and LEA capacity will influence how SEAs attend to these suggestions. We do not expect that all SEAs will create or compile the same set of tools, resources, and/or support, or provide the same set of supporting elements to all LEAs, but the five categories in Figure 7-5 should be considered by all SEAs as high-leverage opportunities to assist LEAs.

As shown in Figure 7-5, we organize the types of tools, resources, and support SEAs can provide into five categories: (1) curriculum and instructional material reviews; (2) a professional learning provider clearinghouse; (3) local assessment practices support; (4) local assessment system auditing tools; and (5) assessment literacy resources. The actions within each category are listed in order of those that require the least to greatest amount of state capacity, involvement, and effort. Although these categories are listed separately in Figure 7-5 and the sections that follow, they are also interrelated.

Curriculum and Instructional Material Reviews

High-quality curriculum and instruction are central to ensuring that all students have access to grade-level, standards-aligned teaching and learning experiences. It is a fundamental equity issue that all students have the opportunity to learn what students statewide are supposed to know and be able to do by the end of each school year. However, the implementation of high-quality curriculum and instructional materials is not occurring in many U.S. classrooms (Kaufman et al., 2020). Instead, many teachers spend inordinate amounts of time supplementing their curriculum for different reasons and with largely unknown effects (Silver, 2022).

High-quality curriculum, instruction, and formative assessment processes—aligned to the content and cognitive complexity of the state's content standards—are the mechanisms by which student learning improves. These factors are especially relevant

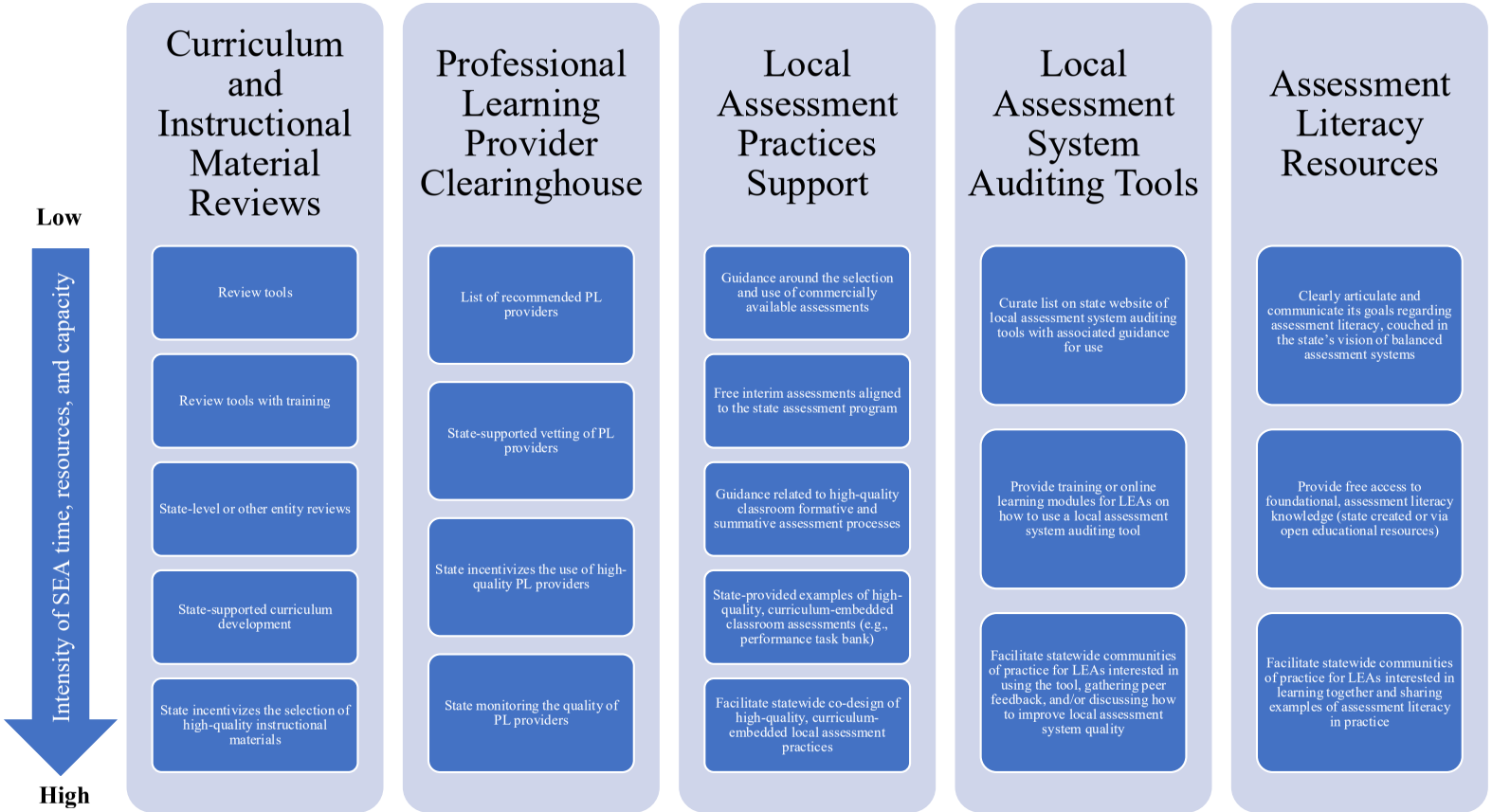


FIGURE 7-5 State tools, resources, and support for LEAs related to balanced assessment systems.

to balanced assessment systems because local assessments should not move students away from high-quality instructional materials—assessments should cohere with and mutually support high-quality teaching and learning experiences within the curriculum. *Ensuring that all LEAs understand what constitutes high-quality instructional materials (HQIMs), including using criteria to evaluate the quality of the curriculum-embedded assessments and assessment processes within those materials, is arguably the most important educational action a state can take to facilitate the design and implementation of balanced assessment systems at the local level.*

Polikoff (2021) convincingly argues that the failure of standards-based reform is due, in large part, to decentralized governance structures that result in poor and inequitable standards implementation. Local control over curriculum results in very little standardization of common curriculum materials across and within states, and there are real differences in curriculum quality that create systemic inequities in students' opportunities to learn. Polikoff argues that radical change is needed to improve instruction at scale and that SEAs have a key role to play in this change process—especially in providing more oversight and support related to HQIMs. One initial action step, Polikoff notes, is for states to collect good data on what curricula are being implemented in classrooms.

Although states have limited control over local curricula, they can provide a variety of related support to facilitate high-quality standards implementation. For example, in 2017, the Council of Chief State School Officers (CCSSO) and a cohort of 12 interested states launched the High-Quality Instructional Materials and Professional Development (IMPD) Network, "dedicated to ensuring that every student, every day, is engaged in meaningful, affirming, grade-level instruction" (Council of Chief State School Officers, n.d.). The IMPD Network provides guidance and case studies to SEAs showing how they can engage with LEAs to adopt HQIMs and ensure access to professional development opportunities that are aligned with those materials (Council of Chief State School Officers, 2021a, 2021b, 2022b, 2022c, 2022d). A recent RAND study on the states in the IMPD Network showed that the incentives the states are using have been effective, especially in mathematics, to create higher rates of adoption and use of standards-aligned HQIMs (Doan et al., 2022).

However, not all states have the same flexibility and capacity. For example, the Wyoming Department of Education operates under a legislative mandate that requires the SEA to remove itself from all local curriculum decisions due to concerns of state overreach. Due to this state statute, the Wyoming Department of Education does not have an office of curriculum and instruction. Therefore, because the provision of state support for HQIMs will be influenced by the size and capacity of the state department of education, as well as a state's legislative freedom around local curriculum, our recommendations below fall along a continuum of state-level involvement.

The five SEA actions that could influence curriculum and instructional material reviews are listed below in order from the least to most required state involvement:

- **Provide high-quality curriculum and instructional material review tools.** The state could focus on signaling the quality of instructional materials to LEAs using HQIMs review tools (Council of Chief State School Officers, 2022b, 2022d). States could also design their own review tools. Additionally, they could adopt or adapt

existing tools, such as EdReports' Curriculum Review Tools (EdReports, 2022b), or state-developed curriculum review rubrics such as those found in Louisiana, Massachusetts, Mississippi, New Mexico, Rhode Island, Tennessee, and Texas (Council of Chief State School Officers, 2022d). It is especially important for states to ensure that any review of curriculum materials interrogates the quality of the curriculum-embedded formative assessment processes and summative assessments to support more balanced assessment systems.

- **Provide training on how to use and apply the review tools.** In addition to providing review tools, SEAs could provide training to help LEA leaders fully understand the tools and practice conducting sample reviews. The training could be self-paced and accessible on demand or provided through in-person workshops. If the state adopts existing review tools, it could promote the corresponding training or certification (EdReports, 2022a). These trainings could also show LEAs potential solutions if they identify gaps in the curriculum—for example, how supplementary materials can be added to achieve more robust curriculum and standards implementation.
- **Conduct state-level review of curriculum and instructional materials (or link to other entity reviews).** Some states directly review curriculum materials. For example, Louisiana has an instructional materials review process where the state provides annotated reviews of K–12 curriculum materials in ELA, math, science, and social studies using evaluation criteria in the state's review tools (Louisiana Department of Education, n.d.). A review produces one of three rankings (Tiers 1–3), reflecting the degree of alignment with the state's content standards and vision of teaching and learning. Although each Louisiana school system can decide whether to draw on these reviews, it is in their best interest to do so insofar as state funding is tied to the selection of Tier 1 curriculum materials. Other states could follow Louisiana's template or decide to follow their own state-level review process, which would make choosing the appropriate curriculum much easier for LEAs. However, the ongoing review of curriculum materials at the state level entails considerable work. Consequently, some states could take advantage of reviews provided by others, such as EdReports (2022b) for English language arts (ELA), math, and science. The state could also consider implementing an EdReports review that is specific to their state (e.g., Arkansas EdReports).
- **Create state curriculum and instructional materials with no adoption requirement.** States could also decide to create their own curricula and offer it to LEAs with no adoption requirement. For example, Louisiana educators have produced K–12 ELA Guidebooks for Louisiana students, which most of the state's school systems use for their ELA curriculum and is offered free of charge to Louisiana school systems. Alternatively, states could partner with an open educational resource curriculum provider to create free, high-quality curriculum materials for their school systems' consideration. For example, 10 states currently partner with OpenSciEd for just this purpose.[6] One advantage of state involvement in curriculum development is that the state can fold its vision for balanced assessment systems directly into curriculum design. For example,

[6] See https://www.openscied.org/why-openscied/partner-states.

the state could create a K–12 curriculum and instructional materials that are accompanied by high-quality formative assessment processes and curriculum-embedded classroom assessments. The practice of developing a tailored curriculum holds considerable promise for disseminating the foundational knowledge necessary to support best practices in classroom assessment.

- **Incentivize the selection of high-quality instructional materials.** The four previous proposed SEA actions focused on states signaling the quality of instructional materials. The final, most time-intensive, but also most impactful action, would be for a state to incentivize the selection of HQIMs. Some states accomplish this by establishing financial incentives for districts that select materials from the state's recommended list of HQIMs. These financial incentives could include state competitive grants, school improvement funding, requirements for use of some federal funds, COVID-19-pandemic-related federal relief funding appropriations, and statewide contracts for HQIMs that reduce the cost of the materials (Council of Chief State School Officers, 2022d).

Professional Learning Provider Clearinghouse

In the same way a state can review curriculum materials, it can also increase the number of teachers who have access to high-quality professional learning about curriculum and standards implementation by incentivizing a strong vendor marketplace. CCSSO's IMPD Network created a guidance document that describes four different ways a SEA can support districts in using high-quality professional learning providers (Council of Chief State School Officers, 2022a):

- provide districts with a list of recommended professional learning providers;
- support districts in vetting professional learning providers;
- incentivize the use of high-quality professional learning providers; and/or
- monitor the quality of professional learning providers.

For example, Louisiana reviews professional learning vendors who target core academic subjects and then provides a vendor guide to all Louisiana school systems.[7] Louisiana then incentivizes the use of high-quality professional learning providers in the Louisiana Super Application, which is an integrated application Louisiana LEAs use to apply for Title I, Title II, and School Improvement funds every year. Additionally, Louisiana has developed a tool to track professional learning provider use and monitors professional learning quality using LEA and Teacher Satisfaction Surveys. The review and dissemination of high-quality professional learning that supports standards implementation, particularly the assessment-related aspects of that training, is an important step in supporting the implementation of balanced assessment systems. The authors recommend reading *Guidance for States on Supporting District Use of High-Quality Professional Learning Providers*, from the Council of Chief State School Officers (2022a) for examples from other states including Delaware and Rhode Island.

[7] See https://www.louisianabelieves.com/academics/curriculum.

As curriculum, instruction, assessment, and state content standards should work coherently together to support balanced assessment systems, these professional learning opportunities should also provide training and coaching to skillfully utilize curriculum-embedded assessments within HQIM and to interpret resulting assessment information. These professional development offerings should include evidence-based implementation features such as ongoing job embedded training and coaching, active learning, teacher collaboration, and involvement of school leaders (Shapovalov & Evans, 2022).

Local Assessment Practices Support

Although states do not control local assessment practices and policies (see Table 7-1), they can indirectly support these activities by providing classroom and interim assessment system tools, resources, and support. High-quality local assessment practices must be in place to support the implementation of balanced assessment systems, as the assessment system is composed mainly of these local assessments. Despite its outsized impact, the state assessment program is a small aspect of a balanced assessment system and has a very particular purpose and intended use. These state-provided tools, resources, and support would be optional, and could include:

- providing guidance around the selection and use of interim assessments;
- providing free interim assessments aligned to the state's theory of action for balanced assessment systems;
- providing guidance related to high-quality classroom formative and summative assessment processes;
- providing examples of high-quality, curriculum-embedded classroom assessments (e.g., performance task bank); and/or
- facilitating statewide support for the co-design of high-quality, curriculum-embedded local assessments.

The first two bullets focus on interim assessments. Interim assessments are optional parts of balanced assessment systems (Marion, 2019a; Marion et al., 2019), although they are ubiquitous and unlikely to fade from use in the near future. Interim assessments are defined as:

> Assessments administered during instruction to evaluate students' knowledge and skills relative to a specific set of academic goals in order to inform policymaker or educator decisions at the classroom, school, or district level. **The specific interim assessment designs are driven by the purposes and intended uses,** but the results of any interim assessment must be reported in a manner allowing aggregation across students, occasions, or concepts. (Perie et al., 2009, p. 6; emphasis added)

The importance of specifying the purposes and intended uses of interim assessments is emphasized in the definition above because interim assessment designs differ and they do not provide the information local educators might need or want equally well (Gong, 2019). There is a strong desire among many educational leaders to procure interim assessments and administer them two to three times over the school year to

gather within-year information on student academic achievement that appears more objective than locally created measures. The state should make decisions on how to support LEAs regarding interim assessments based on its theory of action for balanced assessment systems and the needs of schools and districts. This support could include providing advice about the pros and cons of various interim assessment designs—whether they are commercially purchased or state-provided. It could also include encouraging LEAs to use a thoughtful procurement process that includes specifying use, identifying desired assessment features, and evaluating the technical quality of the interim assessment options (Landl & Lyons, 2023). For example, many LEAs look to commercial interim assessments with the desire to "inform instruction" throughout the year but end up purchasing assessments that are designed to closely mimic the state test design (e.g., NWEA MAP, Renaissance STAR). Unfortunately, local leaders may fail to realize that these types of tests "typically lack sufficient ties to curriculum and instruction to make it possible to provide feedback that leads to improvement" (Shepard, 2005, pp. 2–3). SEAs can provide guidance around interim assessments that helps cut through confusion and marketing claims. SEAs can also decide if providing free interim assessments designed to support specific purposes is important for supporting their theory of action around balanced assessment systems (or not).

The last three bullets above relate to support for high-quality classroom assessment practices. Given rich learning environments foster changes in interactions among the teacher, students, and content, states may want to support the conditions for improved classroom instruction and assessment practices. Classroom assessment tools, resources, and support can range from guidance around best practices to state-provided examples of curriculum-embedded classroom assessments to facilitating statewide gatherings of educators to co-design classroom assessments. These example actions are not mutually exclusive, as a state could support all or only one. For example, the Hawai'i Department of Education is using two recent Competitive Grants for State Assessment awards to design and implement state-provided, optional classroom assessment tools, resources, and support (U.S. Department of Education, 2022). These optional resources focus on increasing the quality of classroom assessments and assessment processes by creating a bank of exemplar performance tasks, tied to the curriculum, with related instructional guides (Hawai'i Performance Assessment Task Bank, 2022). These activities and outputs are intended to build educator capacity to create rich learning environments and ultimately advance student learning through improved classroom instruction and assessment practices.

Local Assessment System Auditing Tools

Part of a state's communication strategy should include sharing its vision of balanced assessment systems with LEAs (see also State Action 2)—but vision sharing is not enough. Rather, LEAs need tools to help them understand what balance means in practice, as well as tools for auditing and evaluating the balance of their local assessment systems (i.e., state-required annual achievement testing, school- and/or district-required assessments, and classroom assessments).

Local assessment system auditing tools and resources can help educators at all levels reflect on the relevance, usefulness, coordination, and quality of the set of assess-

ments that their local assessment system comprises, including state, district and school, and classroom assessments (Chappuis et al., 2016). Assessment audits can help educational leaders identify who needs assessment information when and for what purpose to evaluate the quality of their local assessment system—eliminating inefficiencies, redundancies, and low-quality assessments in the process.

States can support the use of local assessment system auditing tools and resources in several ways:

- curate list on state website of local assessment system auditing tools with associated guidance for use;
- provide training or online learning modules for LEAs on how to use a local assessment system auditing tool; and/or
- facilitate statewide communities of practice for LEAs interested in using the auditing tool, gathering peer feedback, and/or discussing how to improve local assessment system quality.

The quality of local assessment systems is critical because previous analysis has shown that a majority of assessment burden and over-testing arises from locally required assessments (Lazarin, 2014). Local assessment system audits can help promote more balanced assessment systems by evaluating and analyzing the assessments administered in the district or school based on intended users and uses of the assessment information. In other words, is the system of assessments providing the necessary information for specific users to make educational decisions that support student learning at the right time and the right level of specificity and relationship to the enacted curriculum?

Auditing resources (Coladarci, 2002; Council of Chief State School Officers, 2015) and auditing tools for local assessment systems are available for use (Achieve, 2014; Chappuis et al., 2016; EducationFirst, n.d.; Evans & Thompson, 2022b, 2022c; Martineau et al., 2018). The Georgia Department of Education, for example, partnered with the Georgia Partnership for Excellence in Education to pilot how to help school districts test "smarter" rather than more often (Georgia Partnership for Excellence in Education, 2022). Most of these auditing tools require broad stakeholder engagement and provide a comprehensive framework for thinking about local assessment system quality. All of these tools involve time and effort—some more so than others. Only some of these tools include audits of classroom-level assessment systems, as well as state-, district- and school-level. In most cases, the quality of each assessment is not a focus of systems-level evaluations but could be a follow-up or concurrent activity.

Auditing tools need not be overly complex or involve all potential stakeholders to be effective. However, auditing tools should reflect the complexity of systems and the range of students who participate in them (e.g., types and frequency of information needed to support students with disabilities or English learners). The interaction among state-, district-, school-, and classroom-level assessments is important because state assessments may provide information that overlaps with the assessment needs at the district or school level, providing duplicative and redundant information. Without analyzing the assessments altogether, those tasked with auditing and evaluating the system would not see the overlaps and redundancies. The complex web of assessments

must be analyzed together to evaluate the extent to which the entire system exhibits the features of balanced assessment systems.

Assessment Literacy Resources

Assessment literacy encompasses the knowledge and skills that educators (both classroom teachers and school and district leaders) need to appropriately utilize assessments to inform educational decisions about student learning (Stiggins, 1991) (see also Chapter 5 of this volume, "Assessment Literacy and Professional Learning"). The SEA's role in supporting educators' assessment literacy falls on a continuum of involvement. At a minimum, a SEA should clearly articulate and communicate its goals to LEAs regarding educator assessment literacy and how those goals are couched in the state's vision of balanced assessment systems.

After this baseline responsibility, SEAs will have differing amounts of personnel and capacity to support assessment literacy initiatives. In light of the various capacities of SEAs, we are not suggesting that every SEA should create its own assessment literacy professional learning resources. Rather, the SEA can link to free resources, such as webinars and self-paced modules, that have been created by other entities. For example, the Michigan Assessment Consortium provides free resources and tools on its website, and the Center for Assessment provides a set of open-access teacher and leader professional learning modules (Evans & Thompson, 2022a).

The purpose behind a SEA providing free access to foundational assessment literacy knowledge is to create and support school conditions that will promote student learning. Assessment, when it is working as intended, provides feedback loops to students and educators that can be used to adjust teaching to the benefit of student learning. The goal should be to establish a common and sufficient level of assessment literacy knowledge and skill for district-, school-, and classroom-level educators so that they can foster best practices in assessment, student learning, and professional collaboration (DeLuca et al., 2019; Xu & Brown, 2016).

The *Standards for Teacher Competence in Educational Assessment of Students* describe what all teachers should know and be able to do related to assessment in their classrooms (American Federation of Teachers et al., 1990). Others have built on these foundational educator assessment literacy standards (e.g., Klinger et al., 2015; Michigan Assessment Consortium, 2016). There are also assessment textbooks for teachers often used in educator preparation programs (Brookhart & Nitko, 2019; McMillan, 2021). Xu and Brown (2016) synthesize that body of work just listed (and more) and delineate assessment literacy foundational knowledge as disciplinary knowledge and pedagogical content knowledge; knowledge of assessment purposes, content, methods, grading, feedback, and peer- and self-assessment; assessment interpretation and communication; and assessment ethics.

SEAs will likely want to identify what various users need to know and be able to do related to assessment as they consider how to support better educational decision making focused on student learning. For example, district and school leaders should possess literacy about assessment commensurate with their respective roles in the educational system. School leaders are more involved in teacher supervision and instructional coaching, so assessment literacy related to classroom assessment processes

is crucial. District leaders are more involved in making decisions about resource allocation based on test scores, purchasing interim assessments, and setting grading policies so assessment literacy related to those topics is important. Teachers, on the other hand, need training and coaching around high-quality formative and summative classroom assessment processes (see Chapter 5 of this volume, "Assessment Literacy and Professional Learning").

Another way a SEA can support assessment literacy is to facilitate statewide convenings or communities of practice for LEAs that are interested in learning together and sharing examples of assessment literacy in practice. Arranging these convenings will require more involvement from the SEA if the SEA will also be serving as a facilitator, but over time the state could train others to take a leadership role or run more informal networks of support.

To sum up State Action 5, as SEAs increasingly integrate their actions related to curriculum and instructional material reviews, a professional learning provider clearinghouse, local assessment practices support, local assessment system auditing tools, and assessment literacy resources, it is also increasingly likely that the SEA's efforts will lead to systemic and scalable change. SEAs and LEAs must work together as partners to accomplish school reform. It may be that the most important role of the SEA is to serve as a convener and collaborator for local leaders and classroom educators to work together toward lasting education reform.

State Action 6: Engage Educator Preparation Programs

States control teacher, principal, and superintendent standards, licensure, and recertification. SEAs also approve educator preparation programs to ensure that the teachers who graduate from these programs are highly qualified and well prepared to serve all students effectively. We are not suggesting that the solution to ensure all incoming teachers are assessment literate is for SEAs to mandate more coursework for teacher and school and district leader preparation. Rather, we are suggesting that SEAs provide guidance to educator preparation programs about how to integrate and embed assessment literacy principles within core coursework.

States could use their role in the educator preparation program approval process to ensure that the programs provide the coursework and clinical training necessary to support the state's vision and theory of action related to balanced assessment systems. In particular, states could ensure that teachers and school and district leaders who graduate from the state's educator preparation programs understand the importance of HQIMs and have the assessment literacy and content knowledge necessary to support coherence among curriculum, assessment, instruction, and the state's content standards. For example, CCSSO's IMPD Network provides guidance for state policies related to educator preparation to support HQIMs implementation (Council of Chief State School Officers, 2020). This guidance recommends that SEAs use their policy levers—statewide teacher competencies, initial and ongoing program approval or accreditation, and licensure and recertification requirements—to encourage educator preparation programs to revise their coursework and clinical training experiences to align with the state's vision and theory of action. Another example of a state-educator preparation program partnership is the HQIMs labs established between the Arkansas

Department of Education and educator preparation programs in the state (Arkansas Division of Elementary and Secondary Education, 2023), which support the design and implementation of balanced assessment systems so long as the preservice training coherently embeds assessment literacy training within the instructional content and methods training.

CONCLUSION

While many of the decisions that impact local assessment system design and associated policies and practices are made at the district, school, and classroom levels, states play a significant role in promoting the design and implementation of balanced assessment systems. The most important role the state can play in promoting balanced assessment systems is to create and support the right structures and conditions for district and school leaders and classroom educators to be able to do their jobs effectively, thereby improving student learning. States control specific aspects of the education system that can be leveraged into actions to support the right structures and conditions to promote balanced assessment systems. These actions stem from a clear, compelling, and coherent theory of action to achieve a balanced assessment system and include strategic communications, proactively addressing and mitigating issues concerning state assessment programs, and providing LEAs with tools, resources, and support needed for design and implementation of balanced assessment systems at the local level.

Specifically, SEAs should model behaviors, create conditions, and incentivize or facilitate actions that support local efforts to identify or develop assessment tools and practices that provide a comprehensive, coherent, and useful profile of information about student achievement and growth to educators and parents. In this chapter we argue that a SEA can have a significant positive impact on assessment practices at all levels of the educational system by focusing on what it can control and where it has the greatest influence:

- the design of the state's summative assessments, content standards, and curriculum frameworks;
- the implementation of policies that influence or mandate the use of state summative assessment results beyond those that are federally required;
- ensuring clear communication about the intended purpose and use of state summative assessments; and
- the development of tools and resources that provide support consistent with stakeholders' needs and intended role in advancing balanced assessment systems.

At the center of these efforts is a clear vision for teaching and learning and a theory of action that clarifies how assessments prioritized at different levels of the educational system should work together to support this vision. The theory of action should clarify the type and range of information that different stakeholders need to support decision making, the role of the SEA and other stakeholders in ensuring the collection and appropriate use of assessment information, and the necessary conditions for stakeholders to fulfill their intended roles. Articulating this vision is critical for ensuring that the state's actions are consistent with its theory of action, but even more importantly,

articulating this vision ensures that districts and schools have a clear understanding of what it means to design and implement a balanced assessment system. How the SEA situates the state summative assessment program relative to this goal can positively impact what happens at the district, school, and classroom level if the vision is clearly communicated and reinforced through the development of tools, guidance, and other resources that support local efforts.

REFERENCES

Achieve. (2014). *Student assessment inventory for school districts*. https://www.achieve.org/assessment inventory.

American Educational Research Association. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher*, *44*(8), 448–452. https://doi.org/10.3102/0013189X15618385.

American Federation of Teachers, National Council on Measurement in Education, and National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. https://buros.org/standards-teacher-competence-educational-assessment-students.

Arkansas Division of Elementary and Secondary Education. (2023). *Arkansas Initiative for Instructional Materials (AIIM)*. https://dese.ade.arkansas.gov/Offices/learning-services/curriculum-support/arkansas-initiative-for-instructional-materials-aiim.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, *36*(5), 258–267. https://doi.org/10.3102/0013189X07306523.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, *42*(2), 231–268. https://doi.org/10.3102/00028312042002231.

Brookhart, S. M., & Nitko, A. J. (2019). *Educational assessment of students* (8th ed.). Pearson.

Buzick, H. M., Casabianca, J. M., & Gholson, M. L. (2023). Personalizing large-scale assessment in practice. *Educational Measurement: Issues and Practice*, *42*(2), 5–11. https://doi.org/10.1111/emip.12551.

California Department of Education. (n.d.). *Curriculum frameworks & instructional materials: Information on all subject areas*. https://www.cde.ca.gov/ci/cr/cf.

Chappuis, S., Commodore, C., & Stiggins, R. (2016). *Balanced assessment systems: Leadership, quality, and the role of classroom assessment*. Corwin.

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research and Perspectives*, *9*(4), 173–206. https://doi.org/10.1080/15366367.2011.626729.

Coladarci, T. (2002). Is it a house … or a pile of bricks? Important features of a local assessment system. *The Phi Delta Kappan*, *83*(10), 772–774.

Common Core State Standards Initiative. (2021). *Key shifts in mathematics*. https://www.thecorestandards.org/other-resources/key-shifts-in-mathematics.

Conley, D. T., & Darling-Hammond, L. (2013). *Creating systems of assessment for deeper learning*. Stanford Center for Opportunity Policy in Education. https://edpolicy.stanford.edu/sites/default/files/publications/creating-systems-assessment-deeper-learning_0.pdf.

Council of Chief State School Officers. (2015). *Comprehensive statewide assessment systems: A framework for the role of the state education agency in improving quality and reducing burden*. https://www.ccsso.org/sites/default/files/2017-12/Comprehensive%20Statewide%20Assessment%20Systems%20-%20A%20Framework%20Final%206-24.pdf.

Council of Chief State School Officers. (2020). *State Policy on Educator Preparation: A Framework for Levers to Consider*. https://ccsso.org/sites/default/files/2022-01/EPP%20Framework.pdf

Council of Chief State School Officers. (2021a). *High quality instructional materials & professional development network case study: Communications & stakeholder engagement*. https://753a0706.flowpaper.com/CCSSOIMPDCaseStudyCommunications/#page=1.

Council of Chief State School Officers. (2021b). *Overview of CCSSO's high quality instructional materials & professional development network: How states can support the adoption and effective use of high-quality standards aligned instructional materials*. https://753a0706.flowpaper.com/CCSSOIMPDCaseStudyOVERVIEWFINAL/#page=1.

Council of Chief State School Officers. (2022a). *Guidance for states on supporting district use of high-quality professional learning providers*. https://drive.google.com/file/d/1FvJdsUnTeab1apY91z9QM6h0BLr0cg1W/view?usp=sharing.

Council of Chief State School Officers. (2022b). *CCSSO's high quality instructional materials & professional development network case study: An IMPD policy roadmap for states*. https://753a0706.flowpaper.com/CCSSOIMPDPolicyRoadmap/#page=1.

Council of Chief State School Officers. (2022c). *High quality instructional materials & professional development network case study: Impact of the CCSSO IMPD network*. https://753a0706.flowpaper.com/CCSSOIMPDCaseStudyImpact/#page=1.

Council of Chief State School Officers. (2022d). *High quality instructional materials & professional development network case study: Signaling & incentivizing quality*. https://753a0706.flowpaper.com/CCSSOIMPDCaseStudyIncentivizing/#page=1.

Council of Chief State School Officers. (n.d.). *CCSSO's high-quality instructional materials and professional development (IMPD) network*. https://learning.ccsso.org/high-quality-instructional-materials.

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

Dadey, N., & Gong, B. (2017). *Using interim assessments in place of summative assessments? Consideration of an ESSA option*. Council of Chief State School Officers. https://ccsso.org/sites/default/files/2017-12/ASR_ESSA_Interim_Considerations-April.pdf.

Dadey, N., & Gong, B. (2021). *Claims and evidence for through year assessments: Definitions, aims, and use cases*. https://www.nciea.org/wp-content/uploads/2022/07/Through-Year-Convening-Session-1-Slides.pdf.

Datnow, A., & Hubbard, L. (2015). Teachers' use of assessment data to inform instruction: Lessons from the past and prospects for the future. *Teachers College Record*, *117*(4), 1–26. https://doi.org/10.1177/016146811511700408.

DeLuca, C., Coombs, A., MacGregor, S., & Rasooli, A. (2019). Toward a differential and situated view of assessment literacy: Studying teachers' responses to classroom assessment scenarios. *Frontiers in Education*, *4*, 1–10. https://doi.org/10.3389/feduc.2019.00094.

Diamond, J. B., & Spillane, J. P. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record*, *106*(6), 1145–1176. https://doi.org/10.1111/j.1467-9620.2004.00375.x.

Doan, S., Kaufman, J. H., Woo, A., Tuma, A. P., Diliberti, M. K., & Lee, S. (2022). *How states are creating conditions for use of high-quality instructional materials in K–12 classrooms: Findings from the 2021 American Instructional Resources Survey*. RAND Corporation. https://doi.org/10.7249/RRA134-13.

EdReports. (2022a). *Become a reviewer*. https://www.edreports.org/process/reviewers.

EdReports. (2022b). *Review tools*. https://www.edreports.org/reports/review-tools.

EducationFirst. (n.d.). *Fewer and better local assessments: A toolkit for educators (playbook)*. https://www.education-first.com/playbook.

Evans, C. M. (2022, June 22). *Can a state test inform instruction? The typical, the improbable, and the possibilities in the messy middle*. National Center for the Improvement of Educational Assessment. https://www.nciea.org/blog/can-a-state-test-inform-instruction.

Evans, C. M., & Marion, S. F. (in press). *Instructionally Useful Assessment*. Routledge.

Evans, C. M., & Thompson, J. (2022a). *Educational assessment 101 for K–12 teachers and leaders (version 2.0)*. National Center for the Improvement of Educational Assessment. https://www.nciea.org/library/classroom-assessment-learning-modules.

Evans, C. M., & Thompson, J. (2022b). *Leader module 9: What is a local assessment system and how can school/district leaders audit their system to ensure it aligns with the features of balanced assessment systems?* National Center for the Improvement of Educational Assessment. https://docs.google.com/document/d/1uCHOeYIc6fELyEJma68VeF38-j7NvG556WGy01OtSkA/edit.

Evans, C. M., & Thompson, J. (2022c). *Teacher module 23: How do you know if you have a high-quality classroom assessment system?* National Center for the Improvement of Educational Assessment. https://docs.google.com/document/d/17rvK7mZkZAVBpji_WPxTweGaFjm0j6oEYTURa7bkvnA/edit.

Faxon-Mills, S., Hamilton, L. S., Rudnick, M., & Stecher, B. M. (2013). *New assessments, better instruction? Designing assessment systems to promote instructional improvement*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR354.html.

Firestone, W. A., Camili, G., Yurecko, M., Monfils, L., & Mayrowetz, D. (2000). State standards & opportunity to learn in New Jersey. *Education Policy Analysis Archives*, *8*(35). https://doi.org/10.14507/epaa.v8n35.2000.

Fuhrman, S. H., & Elmore, R. F. (Eds.). (2004). *Redesigning accountability systems for education*. Teachers College Press.

Georgia Partnership for Excellence in Education. (2022). *Assessment inventory project*. https://gpee.org/programs/testing-inventory-project.

Gong, B. (2010). *Using balanced assessment systems to improve student learning and school capacity: An introduction*. Council of Chief State School Officers and Renaissance Learning.

Gong, B. (2019). *Matching instructional uses with interim assessment designs*. https://www.nciea.org/wp-content/uploads/2021/11/Gong-Matching-Instructional-Uses-with-Assessment-Designs-100719.pdf.

Hawai'i Performance Assessment Task Bank. (2022). *Hawai'i performance assessment development initiative: Performance assessment task bank & administration guidance.* https://sites.google.com/view/hawaiitaskbank/home.

Kaufman, J. H., Tosh, K., & Mattox, T. (2020). *Are U.S. teachers using high-quality instructional materials?* RAND Corporation. https://doi.org/https://doi.org/10.7249/RR2575.11-1.

Klinger, D. A., McDivitt, P. R., Howard, B. B., Munoz, M. A., Rogers, W. T., & Wylie, E. C. (2015). *The classroom assessment standards for PreK–12 teachers*. Kindle Direct Press. https://evaluationstandards.org/classroom.

Landl, E., & Lyons, S. (2023). *Choosing the right tests: The District Assessment Procurement Protocol (DAPP)*. EdReports.org, Center for Assessment, and Lyons Assessment Consulting. https://www.lyonsassessmentconsulting.com/assets/files/District_Assessment_Procurement_Protocol_FINAL-1.pdf.

Lazarin, M. (2014, October 16). *Testing overload in America's schools*. Center for American Progress. https://www.americanprogress.org/article/testing-overload-in-americas-schools.

Louisiana Department of Education. (n.d.). *Online instructional materials review*. https://www.louisianabelieves.com/academics/ONLINE-INSTRUCTIONAL-MATERIALS-REVIEWS.

Marion, S. F. (2018). The opportunities and challenges of a systems approach to assessment. *Educational Measurement: Issues and Practice*, *37*(1), 45–48. https://doi.org/10.1111/emip.12193.

Marion, S. F. (2019a, October 30). *Do interim assessments have a role in balanced systems of assessment?* National Center for the Improvement of Educational Assessment. https://www.nciea.org/blog/assessment/do-interim-assessments-have-role-balanced-systems-assessment.

Marion, S. F. (2019b, March 21). *Five essential features of assessment for learning*. National Center for the Improvement of Educational Assessment. https://www.nciea.org/blog/innovative-assessment/five-essential-features-assessment-learning.

Marion, S., Thompson, J., Evans, C., Martineau, J., & Dadey, N. (2019). *The challenges and opportunities of balanced systems of assessment: A policy brief*. National Center for the Improvement of Educational Assessment. https://files.eric.ed.gov/fulltext/ED598421.pdf.

Martineau, J., Dewsbury-White, K., Roeber, E., Vorenkamp, E., Snead, S., & Flukes, J. (2018). *District assessment system design toolkit*. National Center for the Improvement of Educational Assessment. https://www.nciea.org/library/district-assessment-system-design-toolkit.

Massachusetts Department of Elementary and Secondary Education. (2023). *Mathematics released items*. http://mcas.pearsonsupport.com/released-items/math.

Mathematics Assessment Resource Service. (2011). *Mathematics assessment project: Balanced assessment: Summative assessment tasks for high school*. https://www.map.mathshell.org/tasks.php?unit=HA07&collection=9.

McMillan, J. H. (2021). *Classroom assessment: Principles and practice that enhance student learning and motivation* (7th ed.). Pearson.

Michigan Assessment Consortium (2016). *Assessment literacy standards*. https://www.michiganassessmentconsortium.org/assessment-literacy-standards.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press. https://doi.org/10.17226/10019.

National Research Council. (2003). *Assessment in support of instruction and learning: Bridging the gap between large-scale and classroom assessment: Workshop report.* The National Academies Press. https://doi.org/10.17226/10802.
National Research Council. (2010). *State assessment systems: Exploring best practices and innovations: Summary of two workshops*. The National Academies Press. https://doi.org/10.17226/13013.
National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. https://doi.org/10.17226/13165.
Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice*, *28*(3), 5–13. https://doi.org/10.1111/j.1745-3992.2009.00149.x.
Polikoff, M. (2021). *Beyond standards: The fragmentation of education governance and the promise of curriculum reform*. Harvard Education Press.
Shapovalov, Y. A., & Evans, C. M. (2022). *Research synthesis: Developing the educator assessment literacy professional learning screener*. National Center for the Improvement of Educational Assessment. https://files.eric.ed.gov/fulltext/ED627088.pdf.
Shepard, L. A. (2005). *Formative assessment: Caveat emptor*. ETS Invitational Conference, The Future of Assessment: Shaping Teaching and Learning. https://csaa.wested.org/wp-content/uploads/2019/11/shepard-formative-assessment-caveat-emptor.pdf.
Shepard, L. A. (2021). Ambitious teaching and equitable assessment: A vision for prioritizing learning, not testing. *American Educator*, *Fall*. https://www.aft.org/ae/fall2021/shepard.
Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, *37*(1), 21–34. https://doi.org/10.1111/emip.12189.
Shultz, P. K., & Englert, K. (2021). Cultural validity as foundational to assessment development: An Indigenous example. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.701973.
Silver, D. (2022). A theoretical framework for studying teachers' curriculum supplementation. *Review of Educational Research*, *92*(3), 455–489. https://doi.org/10.3102/00346543211063930.
State of Connecticut. (2023). *About: Center for Connecticut Education Research Collaboration*. https://portal.ct.gov/ccerc/About-Us.
Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, *72*, 534–539.
Stiggins, R. (2006). *Balanced assessment systems: Redefining excellence in assessment*. Educational Testing Service. https://media.ride.ri.gov/PD/FA/Formative_Assessment_Module_1_Lesson_4/story_content/external_files/Stiggins_Balanced_Assessment_Systems%202006.pdf.
Stiggins, R. (2008). *Assessment manifesto: A call for the development of balanced assessment systems*. ETS Assessment Training Institute. https://famemichigan.org/wp-content/uploads/2018/06/Stiggins-Assessment-Manifesto-A-Call-for-the-Development-of-Balanced-Assessment-Systems.pdf.
Stiggins, R. (2017). *The perfect assessment system*. ASCD.
Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, *10*, 211–227. https://doi.org/10.1007/s10833-009-9105-2.
The Leadership Conference on Civil and Human Rights. (2015). *Civil rights groups: "We oppose anti-testing efforts."* https://civilrights.org/2015/05/05/civil-rights-groups-we-oppose-anti-testing-efforts.
The Office of Governor Ned Lamont. (2023, March 22). *Governor Lamont announces plan to launch intensive mathematics tutoring program for students in grades 6–9 to address learning loss from pandemic*. Press release. https://portal.ct.gov/Office-of-the-Governor/News/Press-Releases/2023/03-2023/Governor-Lamont-Announces-Plans-to-Launch-Intensive-Mathematics-Tutoring-Program.
U.S. Department of Education. (2022). *Awards*. https://oese.ed.gov/offices/office-of-formula-grants/school-support-and-accountability/competitive-grants-for-state-assessments/awards.
Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010.