



CREATING A STATE VISION

to Support the Design and
Implementation of an Innovative
Assessment and Accountability System

July 2016



Acknowledgements







Thanks to generous support from the Nellie Mae Education Foundation, KnowledgeWorks and the National Center for the Improvement of Educational Assessment (Center for Assessment) have partnered to help states better understand and leverage the new Innovative Assessment and Accountability Demonstration Authority authorized under the Every Student Succeeds Act (ESSA). The goal of this partnership is to help states identify and explore a set of readiness conditions that are critical to the development of a high quality application and implementation process under this new authority. While we share a history of advocacy for next generation assessments, our organizations each bring a unique perspective to this work. KnowledgeWorks focuses on policy development, partnering with states, districts, and educators to identify and remove policy barriers that inhibit the growth of personalized learning. The Center for Assessment specializes in the design of assessment and accountability systems, helping states, districts, and other entities improve the quality of these systems and maximize student success.

Table of Contents

Introduction	4
<hr/>	
Overview	5
Innovative Assessment System Defined	5
Federal Flexibility	5
Guardrails	6
Demonstration Authority Timeline	7
Why would a State Need to Apply for the Demonstration Authority	7
Decision Tree	8
Purposes and Unique Characteristics of the Demonstration Authority	9
State Example	9
<hr/>	
State and District Roles	10
Consideration for State and District Leaders	10
<hr/>	
A Theory of Action to Support the Design and Evaluation of the Innovative Pilot	11
General Design Consideration	11
Getting Started: Eight Essential Steps	12
Sample Theory of Action	14
<hr/>	
Cross-cutting Elements	15
<hr/>	
Summary	16
<hr/>	
Additional Support	17
<hr/>	
About	18
KnowledgeWorks	18
National Center for the Improvement of Educational Assessment	18
Nellie Mae Education Foundation	18

Introduction

This is the first in a series of policy and practice briefs produced by KnowledgeWorks and the National Center for the Improvement of Educational Assessment (Center for Assessment) designed to assist states in thinking through the opportunities and challenges associated with flexibility provided under the Every Student Succeeds Act (ESSA). These briefs help define “Readiness Conditions” for states considering applying for and successfully implementing an innovative assessment and accountability system as defined by the Demonstration Authority opportunity under ESSA. The following briefs will be released over the next few months:

-  **Ensuring and Evaluating Assessment Quality**
-  **Addressing Accountability Issues including Comparability**
-  **Professional Learning and Developing Assessment Literacy**
-  **Constructing a Research and Evaluation Plan**
-  **Building Capacity and Resources**
-  **Scaling Statewide and Ensuring Long Term Sustainability**

Overview

The newly-enacted Every Student Succeeds Act (ESSA) provides states with a significant opportunity to develop an innovative assessment system that supports the state’s vision for student-centered, personalized learning or other systems designed to promote deeper and more engaged learning. While there are a number of provisions in ESSA that states can leverage to build these systems, the Innovative Assessment and Accountability Demonstration Authority (hereafter known as the “innovative pilot” or the “Demonstration Authority”) authorized under Section 1204 provides states with an unprecedented opportunity to develop next generation approaches to assessment that transcend the standardized instruments commonly used to evaluate student and school performance.

ESSA Sec. 1204—Innovative Assessment System Defined

An Innovative Assessment System means a system of assessments that may include:

- (1) competency-based assessments, instructionally embedded assessments, interim assessments, cumulative year end assessments, or performance-based assessments that combine into an annual summative determination for a student, which may be administered through computer adaptive assessments;
- (2) assessments that validate when students are ready to demonstrate mastery or proficiency and allow for differentiated student support based on individual learning needs.

Federal Flexibility

When Congress established the Demonstration Authority, it did not intend for the program to serve as a rubber stamp for one particular assessment design. Instead, Congress hoped to support states in the creation of new, innovative approaches to assessment and accountability that provide better information about student learning so stakeholders can make continuous improvements to the education system. States interested in this authority have the opportunity to apply to the U.S. Department of Education (ED) for flexibility from several federal assessment requirements. States will need to submit a high quality plan that outlines the proposed assessment approach along with a compelling theory of action for how the system will support comprehensive education reform statewide.

The Demonstration Authority will provide states an opportunity to seek the following flexibilities:

- **Assessments are not Required to be the Same Statewide**—Approved states may pilot the assessment system with a subset of districts before scaling the system statewide by the end of the Demonstration Authority. Without this authority, states are required to administer the same academic assessments to all students in the state.
- **Assessments May Consist Entirely of Performance Tasks**—Approved states may design an assessment or system of assessments that consists of all performance tasks, portfolios, or extended learning tasks. Without this authority, states are only permitted to use performance tasks for part of the assessment.
- **Assessments May Be Administered When Students Are Ready**—Approved states may assess students when they are ready to demonstrate mastery of standards and competencies as applicable. Without this authority, states may assess students using a single, summative instrument or as a combination of interim assessments that in aggregate provide one summative score.

Guardrails

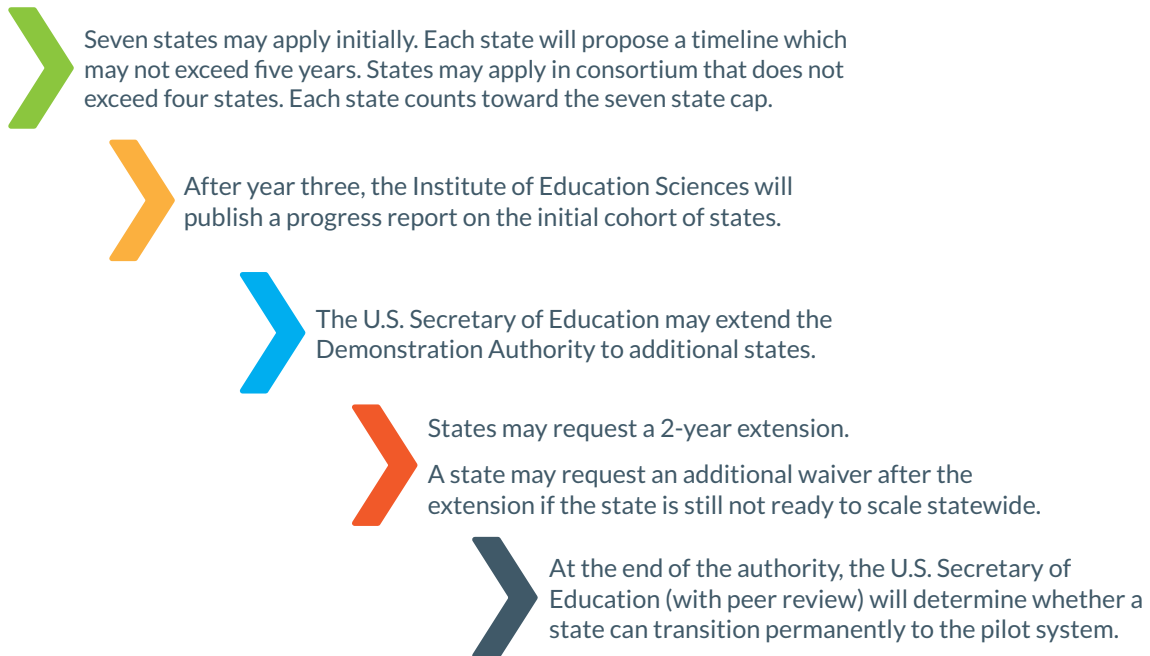
To ensure high quality implementation across pilot sites, the Demonstration Authority includes a set of guardrails that participating states will have to meet throughout the demonstration process. While all participating states will be held to the same guardrails, the evidence that they submit to the ED to demonstrate adherence to these provisions will be different depending on the state's assessment design and theory of action.

The Demonstration Authority includes the following guardrails:

- **Assessment Quality**—"The state needs to demonstrate that the system of assessments is comprised of high quality assessments that support the calculation of valid, reliable, and comparable annual determinations as well as provide useful information to relevant stakeholders about what students know and can do relative to the learning targets."
- **Comparability**—"The state needs to demonstrate that its innovative assessment system produces yearly, student-level annual determinations that are comparable across local education agencies (LEAs) and to the federally required statewide assessments and for each subgroup of students as compared to the results for such students on federally required state assessments."
- **Scale Statewide**—"If the state is proposing to administer the innovative assessment system initially in a subset of LEAs, the state must have a logical plan to scale up the innovative assessment system statewide in the State's proposed demonstration authority period."
- **Demographic Diversity & Similarity**—"The state can describe how the inclusion of additional LEAs will help the state make progress toward achieving high-quality and consistent implementation across demographically diverse LEAs. The state can also describe how it will ensure that the participating LEAs, as a group, will be demographically similar to the state as a whole by the end of the state's demonstration authority period."

Demonstration Authority Timeline

The timeline for the Demonstration Authority begins when the U.S. Secretary of Education issues an application to states. ESSA gives the Secretary full discretion to decide if and when to begin the innovative pilot.



Why Would a State Need to Apply for the Demonstration Authority?

A state should consider pursuing the Demonstration Authority if it is more likely to accomplish the education goals aligned to its vision with one or more of these policy flexibilities. For example, a state interested in establishing a competency-based learning system where students are assessed when they are ready to demonstrate achievement would need flexibility from the time-bound ESSA requirement that states administer an end-of-year standardized summative assessment at the same time for all students. However, designing, implementing, and maintaining such an innovative system will require significant work, so states will need to weigh the “costs” compared with the opportunities of implementing such a system. States also have more flexibility for assessment design and selection under the general assessment requirements of ESSA (Section 1111) than they did under the No Child Left Behind (NCLB) Act. The Council of Chief State School Officers (CCSSO) issued a decision tree guide to help states explore these opportunities and determine whether they need the Demonstration Authority to meet their assessment needs¹. Figure 1 includes one of the decision frameworks from this guide.

A state will need to determine early in the design phase whether it intends to engage in its reform plan with a subset of districts or to move to statewide adoption from the onset. Depending on the scope of the innovation, a state that is interested in beginning with statewide adoption may decide it does not need to pursue the Section 1204 innovative pilot because it may have the flexibility necessary to pursue its agenda under Section 1111.

¹ [http://www.ccsso.org/Documents/2016/CriticalAreaOutlineInnovativeAssessmentPilots\(0\).pdf](http://www.ccsso.org/Documents/2016/CriticalAreaOutlineInnovativeAssessmentPilots(0).pdf).

DECISION TREE

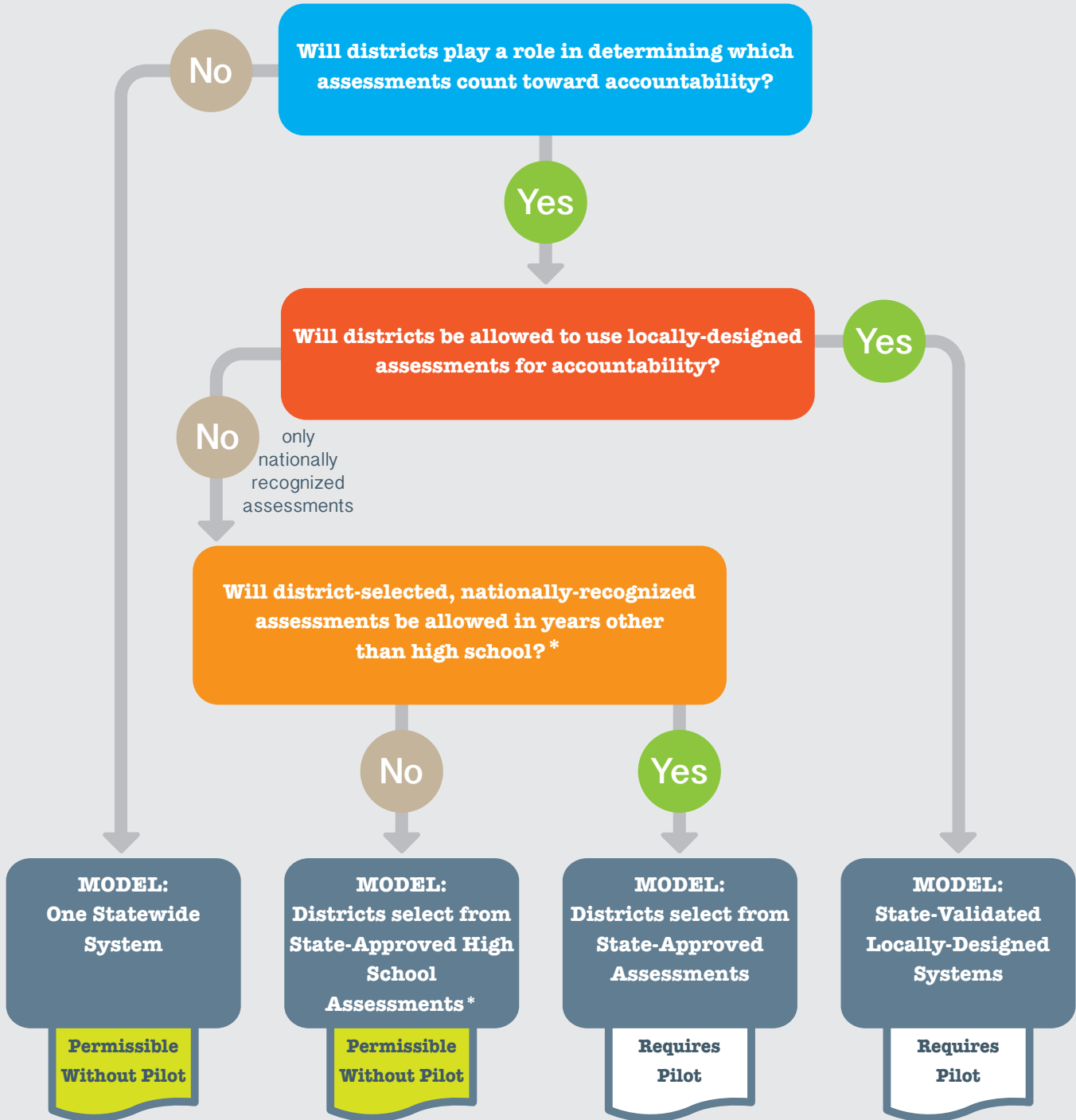


Figure1. Decision tree for determining whether a state needs to apply for an innovative pilot to take advantage of assessment flexibility (reprinted with permission from the Council of Chief State School Officers (CCSSO)).

Purposes and Unique Characteristics of the Demonstration Authority: It's All About the System!

States should not pursue the Demonstration Authority if their goal is simply to reduce the use of summative tests in grades 3–8 and once in high school. Rather, the decision to pursue an innovative pilot should be tied to a larger reform effort that calls for important changes to the teaching and learning process.

While the law specifically references “competency-based assessment” in its definition of an innovative assessment system, systems of assessment designed to support competency-based and personalized learning environments are not the only reform initiatives possible under the Demonstration Authority. For example, state leaders may want to support local assessment capacity through the implementation of balanced assessment systems that will lead to greater educator agency and ultimately deeper learning by students. Whatever vision the state develops, it must articulate how the flexibility afforded by the Demonstration Authority will help the state achieve defensible goals to improve student learning and increase the productivity of the education system.

Further, states exploring the Demonstration Authority must remember that the assessment system will operate within a larger sociopolitical context. A key piece of this larger context is the federally-required school accountability system. When Congress created the Demonstration Authority in ESSA, policymakers intentionally gave it the name: The Innovative Assessment and Accountability Demonstration Authority. States will need to consider how the participating subset of districts (assuming the pilot starts with a subset) will generate high-quality information about student learning to inform state accountability determinations and, how the state will reorient and enhance accountability for the pilot districts in ways that are coherent with the proposed innovative learning and assessment systems.

STATE EXAMPLE

New Hampshire, operating under a waiver from the U.S. Department of Education, similar in nature to the Demonstration Authority, has reoriented and enhanced its accountability system for participating districts. New Hampshire’s Performance Assessment of Competency Education (PACE) is a competency-based educational approach designed to ensure that students have meaningful opportunities to achieve critical knowledge and skills. Richard Elmore’s concept of reciprocal accountability is at the core of New Hampshire’s model:

For every increment of performance I demand from you, I have an equal responsibility to provide you with the capacity to meet that expectation. Likewise, for every investment you make in my skill and knowledge, I have a reciprocal responsibility to demonstrate some new increment in performance (Elmore, 2002, p.5).

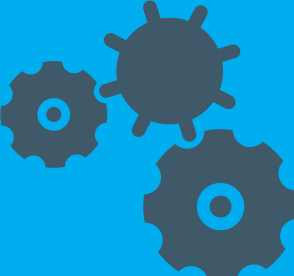
For PACE, this means local educational leaders are involved in designing and implementing the assessment and accountability systems and receive intense technical, policy, and practical support and guidance from the state department of education. The state agency must still be able to evaluate student and school performance using the same accountability indicators that all other districts in the state follow, however, accountability is enhanced under PACE because the locus of accountability has shifted to a more internal orientation compared to a top-down approach. Again, states interested in the Demonstration Authority do not have to adopt a reciprocal accountability framework, but they will need to be explicit in how they intend to connect their learning, assessment, and accountability systems to further their educational goals while supplying the required information for the state’s ESSA accountability system.²

²Our forthcoming brief on accountability and comparability will explore these issues in greater detail.

State and District Roles

Whether the state adopts a reciprocal accountability framework or another approach, successfully implementing an innovative assessment and accountability system requires a strong collaborative working relationship among the participating districts and the state agency responsible for the pilot. Most innovative pilot initiatives will require more work for local districts than simply maintaining the current system. We argue that it is the “right work,” but it is still additional work compared to what districts are currently doing to support and administer the state assessment and accountability system.

Considerations For Leaders

District	State
<p>As full partners in the innovative pilot, district leaders will need to:</p> <ul style="list-style-type: none">• Commit to the shared educational vision associated with the innovative pilot and ensure that the local school boards, educators, and other relevant stakeholders understand the expectations associated with the pilot;• Invest in building capacity among the staff to implement the reforms in instruction and assessment;• Structure or restructure budgets to provide funding for professional development, planning time, assessment development and field testing, data collection, and other considerations;• Create time for collaborative and individual planning, including creating student and teacher schedules to support the learning innovations;• Implement record keeping and student management systems designed to support the innovative assessment system, which may require the capacity to upload student work samples and other data collection needs; and• Most importantly, district leaders and district educators who agree to participate in the pilot must embrace the notion that increased flexibility brings with it increased responsibility. This is the real movement towards an internally-oriented accountability system.	<p>State leaders on the other hand, will have to learn to give up the control that is generally associated with a top-down approach to assessment and accountability. Instead, state departments of education will have to:</p> <ul style="list-style-type: none">• Shift to a more support-oriented organization for the pilot to succeed, including:<ul style="list-style-type: none">– collaborating with participating districts to establish the goals of the pilot, and– designing the most effective ways of reaching those goals through involvement in creating, and carrying out key functions associated with, a theory of action (described in detail below);• Lead the fundraising efforts because many innovative pilot designs will require external funding, at least for the initial years in order to have a chance of success; and• Serve as the liaison with ED, providing evidence that the pilot is meeting the promised technical requirements. 

A Theory of Action to Support the Design and Evaluation of the Innovative Pilot

A “theory of action” is a critical tool for the design and evaluation of policy. It can help illuminate alternatives and competing or even contradictory claims about how a policy initiative should work. Policymakers and designers must explicitly lay out how the proposed design choices are intended to accomplish the goals of the program. In other words, why is the pilot designed in the way it is? In addition to the why, policymakers must also describe the how. Having to articulate both the aims and mechanisms of the program will expose proposed policies for evaluating schools that may be untenable and will also shed light on some fruitful means of meeting the major policy goals.

As states design and implement innovative assessment and accountability systems, they must explicitly state the crucial link between the purposes and uses of the assessment results, and how those results will lead to what the innovative pilot is ultimately designed to achieve. When outlining the theory of action, the design team needs to ensure that connections among various aspects of the assessment system are not simply belief statements, but they can be supported by research, ideally, or at least best practice if research is not available. Given the innovative nature of the assessment and accountability systems likely to be designed under the Demonstration Authority, the links between the elements of the theory of action should additionally serve as testable hypotheses that can be verified with evidence through the implementation of the pilot. This accumulation of evidence would support the ongoing validation of the assessment and accountability systems.

General Design Considerations

Some general considerations for designing a theory of action for an innovative pilot include a description of how:

- a. Each component of the proposed innovative pilot will be clearly and convincingly related to the other components in the system (e.g., how the components will work together to achieve the desired outcome);
- b. The assessment results and accountability indicators will be used;
- c. The assessment results will be incorporated into a coherent educational system (i.e., a system that includes standards, assessments, curriculum, instruction, and professional development);
- d. The innovative pilot as a whole will improve student achievement and college- and career-readiness (or other relevant goals);
- e. The capacity of participating educators and organizations will improve; and
- f. The system will be expanded to include all districts in the state.

One of the benefits of designing a policy initiative by starting with a theory of action is that potential unintended, negative consequences may be identified and mitigated ahead of time by checking the systemic assumptions that must hold in order for the system to function as intended. This check on the logic of the underlying assumptions of the various proposals will serve as an important touchstone during the design process. Again, a theory of action is not just a bunch of pretty shapes and arrows. It must be an empirically and logically based argument that outlines how the specific proposed system will fulfill the stated goals.

Getting Started: Eight Essential Steps

There is no single approach for creating a theory of action, but states may find the following steps useful.

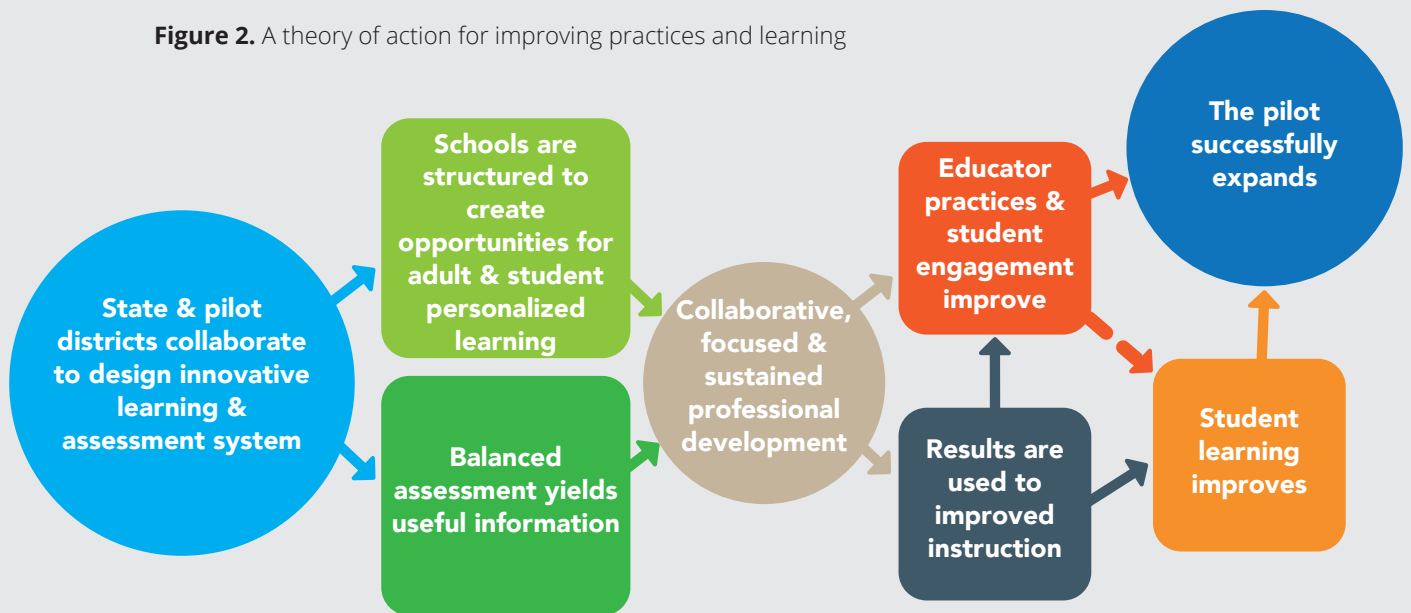
- 1** — Clearly describe the **goals** of the innovative pilot. It is quite likely that there will be multiple goals for the system, but the state and participating districts should try to narrow these down to the highest priority and highest consensus goals. These goals will certainly include the specific outcomes (see step #3), but will also include broader goals for the educational system.
- 2** — The next step is to articulate the **purposes** and intended **uses** of the assessment system results. Being as clear as possible about the goals, purposes, and intended uses up front helps provide the foundation for the theory of action. (The purposes and uses of the system will be linked directly to attainment of the goals of the system through steps 3–7.)
- 3** — The state and pilot districts should come to a rough agreement on the specific intended **outcomes** of the system. For example, a likely intended outcome for most innovative pilots will be to increase student engagement and ultimately the rates of college and career readiness (CCR) for all students. This outcome or outcomes will be closely related to the goals for the system.
- 4** — The next step is to start laying out the **mediating outcomes** necessary to achieve the ultimate outcome(s). Using the example of improving the rates of CCR for all students, some important mediating variables could include such things as: a) “teachers will engage students in meaningful learning activities;” b) “students will learn to direct or at least co-direct their own learning;” c) “student growth trajectories will improve over time;” and d) “teachers (after receiving useful assessment information) will improve their instruction and learning activities.” These are just a few of many examples and the reader should note that some of these would be influenced by prior mediating outcomes and each would need to be expanded by clarifying the mechanisms (see step #7).
- 5** — We have found it helpful to create an initial “high-level” (large grain size) theory of action as a first step in pulling together the results of steps 1–4. This lays out the big picture components and illustrates how these major components are intended to relate to one another.

- 6 — Once the high level theory of action is created, design teams should add enough details to articulate how these major components relate to the minor components. At a certain point, the design team will need to decide the level of detail that can be represented in a single diagram, if this is done pictorially, or in a single set of written steps or statements.
- 7 — The final step involves “zooming in” on several key components of the theory of action to add the detail necessary to support the innovative pilot design and the validity argument. This step is crucial because this is where the design teams have the opportunity to specify the hypothesized mechanisms by which the intended intermediate and final outcomes are thought to occur. Through these mechanisms, the **uses** of the innovative assessment results will be linked to attainment of the **goals**. For example, a theory of action might suggest that providing information at the competency-level will lead to improved student learning. In this case, the state/district design team, when working at this detailed level, should be expected to hypothesize the mechanisms or processes by which the data from the innovative assessment system will lead to better learning outcomes for students such as the development of intervention programs for students who struggle to attain particular competencies. The specification of this hypothesized mechanism then becomes a claim to be included in the validity argument.
- 8 — Once the chain of logic for attaining system goals is clearly specified, the underlying assumptions which must hold in order for the system to function as intended should be articulated. To continue the example from step #7, if reporting competency-based assessment scores is intended to lead to improved student achievement through the use of targeted interventions, the assumption that the intervention is effective must be upheld. Assumptions such as this must be clearly stated in order to identify the conditions under which the goals are most likely to be attained. Through this process of articulating assumptions, we will find that some assumptions are more likely to be violated than others. In the event we identify an assumption that is either likely to be violated, or if violated, consequences would be dire, this will signal a need to potentially revisit the design of the system and revise the theory of action.

Sample Theory of Action

We present a brief example of a high-level theory of action for an innovative pilot system shown in Figure 2 below. In reality, a theory of action used in practice would have many more details. As seen in the figure below, the “focused and sustained professional development” is the mediating mechanism through which teachers get feedback on their practices and learn how to translate assessment information into useable instructional strategies. The assumptions that must hold for this to be true would be added alongside the connecting arrows.

Figure 2. A theory of action for improving practices and learning



Cross-cutting Elements

As a state develops its goals and theory of action we recommend paying special attention to three overarching, cross-cutting elements that will drive a high quality education system and coherently connect the design of the innovative assessment system to the proposed learning systems:

- **Equity**—As innovative approaches to education take root as a result of rethinking current systems, all students must have access to rigorous content and high-quality instruction that meets their individual needs. States and districts should work together in the innovative pilot design phase to carefully consider the supports and interventions that must be in place to ensure that schools have the resources and expertise necessary to close achievement gaps.
- **Continuous Improvement**—The process of implementing new, innovative and potentially personalized approaches to education will result in both successes and challenges. At all levels, the system must adopt data-driven strategies to evaluate system performance and make intentional adjustments and refinements to maximize outcomes.
- **Transparency**—As the state seeks to design, connect, and implement its new learning, assessment, and accountability systems, engaging stakeholders at all levels is essential to successful design and implementation. Transparently engaging key stakeholders such as educators, district leaders, school boards, parents, and other education partners is critical not only to buy-in and understanding but also to gain insights from the field that are critical in a successful system.

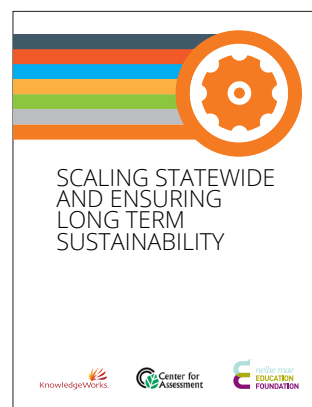
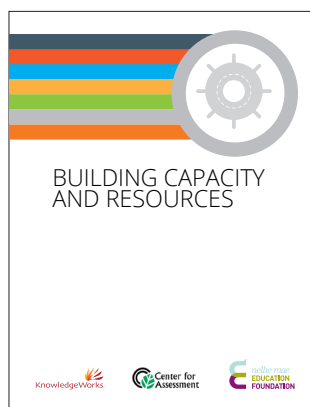
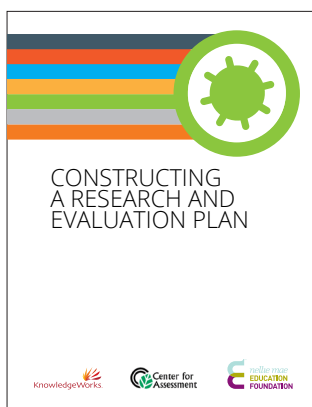
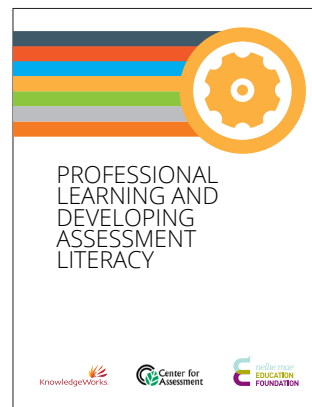
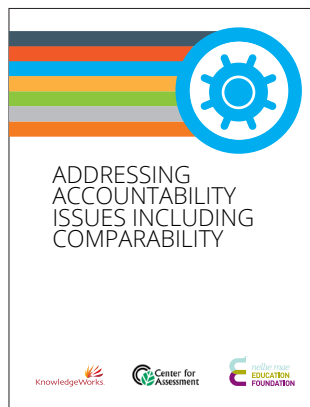
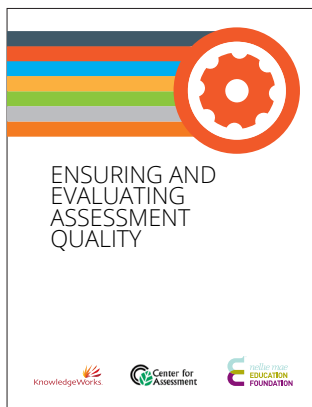
It is important to note that these three overarching, cross cutting elements are critically important not only in the development of a theory of action but also throughout the process of developing an innovative assessment and accountability system. These elements will be carried forward as a consistent thread throughout the series of policy and practice briefs.

We intend for this discussion to serve as a useful starting point for states considering applying for an innovative pilot. While the steps outlined above are just one of many possible approaches for generating a theory of action for an innovative assessment and accountability system, we hope this brief offers a useful process for getting started with this important task.

Summary

This first brief in our series of State Readiness Conditions publications is designed to help state leaders recognize the opportunities associated with the Innovative Assessment and Accountability Demonstration Authority flexibility afforded under the recently passed ESSA. We also present a realistic picture of the challenge associated with the design and implementation of an innovative pilot. This brief help states clarify the need and desire for pursuing an innovative pilot application and explain how developing a theory of action will be a crucial planning and evaluation tool for guiding the design of an application and innovative assessment and accountability system.

KnowledgeWorks and the Center for Assessment will continue to support states through the summer and fall with additional briefs on topics related to fleshing out the design of a Demonstration Authority application, including:



Additional Support

KnowledgeWorks and the Center for Assessment are available to help states as they explore, design, and implement next generation assessment systems. Contact information for our organizations is listed below.

KnowledgeWorks can help states, districts, and other interested stakeholders establish the policy environments to support personalized learning at scale. The organization's expertise spans the federal, state, and district levels, supporting states with strategies to leverage current policy opportunities, remove existing policy barriers, and develop new policies that will help states create an aligned policy environment to support personalized learning. To learn more, contact the following people:

For State Policy and Alignment:

Matt Williams
Vice President of Policy and Advocacy
Williamsm@knowledgeWorks.org

For Federal Policy and Alignment:

Lillian Pace
Senior Director of National Policy
pacel@knowledgeWorks.org

The Center for Assessment strives to increase student learning through more meaningful educational assessment and accountability practices. We engage in deep partnerships with state and district education leaders to design, implement, and evaluate assessment and accountability policies and programs. We strive to design technically sound policy solutions to support important educational goals. The Center for Assessment's professionals have deep expertise in educational measurement, assessment, and accountability and have applied this expertise to assessment challenges ranging from improving the quality of classroom assessments to ensuring the technical quality of state's large-scale achievement tests and ultimately to designing coherent assessment and accountability systems.

For Assessment and Accountability System Design and Strategic Implementation:

Scott Marion, Ph.D.
Executive Director
smarion@nceia.org

For Technical Quality and Comparability Design and Analyses:

Susan Lyons, Ph.D.
Associate
slyons@nceia.org

For Assessment Quality and Performance Assessment Development:

Jeri Thompson, Ed.D.
Senior Associate
jthompson@nceia.org

About Us



KnowledgeWorks is focused on ensuring that every student experiences meaningful personalized learning that allows them to thrive in college, career and civic life. We develop the capabilities of educators to implement and sustain competency-based and early college schools, work with state and federal leaders to establish aligned policy conditions and provide national thought leadership around the future of learning. Our subsidiary, StriveTogether, manages the national cradle-to-career network and helps communities to improve student outcomes through collective impact. www.knowledgeworks.org



The National Center for the Improvement of Educational Assessment, Inc. (Center for Assessment) is a Dover, NH based not-for-profit (501(c)(3)) corporation that seeks to improve the educational achievement of students by promoting enhanced practices in educational assessment and accountability. The Center for Assessment does this by providing services directly to states, school districts, and other organizations regarding the design, implementation, and evaluation of assessment and accountability systems. As a non-profit organization committed to the improvement of student learning, the Center for Assessment maintains a strong “open-source” ethic in terms of distributing its many creations and inventions. For example, the Center has developed many tools related to alignment methodology, student growth analyses, student learning objectives, comparability methods for innovative assessment systems, and validity evaluation that it provides freely to its clients and other non-commercial entities. www.nciea.org



The Nellie Mae Education Foundation is the largest philanthropic organization in New England that focuses exclusively on education. The Foundation supports the promotion and integration of student-centered approaches to learning at the middle and high school levels across New England—where learning is personalized; learning is competency-based; learning takes place anytime, anywhere; and students exert ownership over their own learning. To elevate student-centered approaches, the Foundation utilizes a four-part strategy that focuses on: building educator ownership, understanding and capacity; advancing quality and rigor of SCL practices; developing effective systems designs; and building public understanding and demand. Since 1998, the Foundation has distributed over \$180 million in grants. For more information about the Nellie Mae Education Foundation, visit www.nmefoundation.org.