Strengthening Claims-based Interpretations and Uses of Local and Large-scale Science Assessment Scores

Theory of Action Development Guide

November 2017

Table of Contents

Purpose

All assessments are designed with a purpose in mind, and only by identifying and clarifying this purpose, or set of purposes, can one begin to determine how to evaluate the validity of the interpretations of the scores an assessment yields. A principled-design approach to assessment development enables state assessment systems to be set up in such a way that demonstrates that the end goals of the system were thought about during the design and development phase. This is achieved, in part, through the development of a Theory of Action (ToA), which demonstrates the claims and assumptions that must hold true to support the interpretation(s) and use(s) of assessment scores. Development of a ToA is essential for states to better articulate how their assessment claims connect with, and are supported by, test scores and other sources of evidence. This deep analysis of a state's argument for score meaning helps to strengthen both the validity and coherence of their system. Such an approach also provides stakeholders with ample documentation of design and development logic and decisions, which can be used for future learning, evaluations, and development projects.

Further, developing a ToA through the implementation of a principled-design approach is a key first step to ensuring that assessment development activities and objectives meet the standards of the professional testing community as communicated through the *Standards for Educational and Psychological Testing* (hereafter referred to as the *Standards*; AERA, APA, & NCME, 2014). The *Standards* are the primary guidelines used to improve upon current practices and develop new processes for assessment system evaluation and design. The ToA is an essential element of an assessment system's design that directly supports Standard 1.0: "Clear articulation of each intended test score interpretation for a specified use should be set forth, and appropriate validity evidence in support of each intended interpretation should be provided" (AERA, APA, & NCME, 2014, p. 23).

Thus, each state in the Strengthening Claims-based Interpretations and uses of Local and Large-scale Science Assessment Scores (SCILLSS) project developed a state-specific ToA for their use, and contributed to the development of a common project ToA that reflects the processes, activities, and desired outcomes shared by the participating states. These ToAs allow the three states and the project to reflect on how the project and state assessment systems function, and will guide states through each of the project activities.

Guiding Questions

The ToA is a logic model for how the entire assessment system must function for scores to provide meaningful evidence for intended purposes and uses, and for students to reach the intended outcomes. For each of the components of the ToA, it is up to the states to articulate the guiding philosophy behind their system(s). Below, we have provided guiding questions for each of the five ToA components:

- **Statewide Assessment System Design:** What are the assessment system claims? How is the assessment system designed? How must the assessment system function to provide interpretable and usable scores?

- **System Setting and Use:** How are stakeholders meant to use assessment information? What are some of the conditions that must be in place for the assessment system to function as intended?

- **Teacher Actions:** What activities are expected of teachers? How do teachers interact with students in the classroom? How do teachers use student work to track progress?

- **Student Actions:** What activities are expected of students? How do students interact with teachers and other students? How do students track their progress?

- **Student Outcomes:** What are the intended student goals, outcomes, or consequences of the assessment system (e.g., for students, teachers, instruction)?

From the purposes, goals, and guiding philosophies represented in the ToA, the specific assessment-related claims or issues that are critical to support score meaning are identified.

Development of the State-specific and SCILLSS Project Theories of Action

Development of the state-specific and common project ToAs occurred both in-person and virtually. At the June 2017 Kick-off Meeting, project staff initiated development of a) the state-specific ToAs, and b) the common project ToA. Project staff then held virtual meetings to conduct additional reviews and completed final drafts of the ToAs after the SCILLSS Kick-off Meeting. As previously mentioned, these ToAs illustrate the claims and assumptions that must hold true to support states in the interpretation and use of science assessment scores.

In-Person Development Activity

Validity evaluation experts convened all project staff and provided a comprehensive overview of principled-design, how ToAs fit within that approach, and the goals of the ToA for both the SCILLSS project and each individual state. The project staff worked in state-specific groups in the same room, with one validity evaluation expert assigned to each group. Across the room, facilitators placed large, blank pieces of paper that represented each of the ToA components. Experts provided state staff with paper, pens, and highlighters to use for brainstorming ideas for each of the ToA components.

State-specific Theories of Action

Together with a group facilitator and a validity evaluation expert, state staff spent three hours on the first day brainstorming ideas for each of the ToA components, taking into consideration their state-specific contexts and how the SCILLSS activities and approach fit within their state activities and goals. The guiding questions provided earlier in this document assisted states in brainstorming ideas for each of the components. As they arrived at ideas for each of the ToA components, the group facilitator captured those ideas on corresponding large, blank pieces of paper. Facilitators utilized the large pieces of paper to support states in identifying their commonalities, which ultimately aided the SCILLSS partners in creating a common project ToA.

Upon completion of the brainstorming activity, each state worked with their facilitator and validity evaluation expert to refine their ideas for each of the components of the state-specific ToA. The facilitator led the state staff in a discussion to reach consensus for each component, assisting to clarify language when needed. Furthermore, the facilitator and validity evaluation expert assisted the state staff in articulating the ToA as both a pictorial representation and in paragraph form. At the end of the day, state partners shared out across the groups their drafted, state-specific ToAs.

SCILLSS Project Theory of Action

During the state report out and during break, the facilitators and validity evaluation experts reviewed the drafted ToAs and identified common themes and assessment claims. The validity evaluation experts brought the themes forward with project staff to begin discussion around the common project ToA. Validity evaluation experts discussed the themes identified, and asked states to discuss any additional

themes appropriate for the project. Project staff used these themes to identify the broader, project-level activities that feed into the ToA. The validity evaluation experts worked with staff in drafting each component, beginning with the Student Outcomes and then starting again from the Statewide Assessment System Design. All project staff reached consensus on the components, which the large group facilitator and validity evaluation experts measured with a raise of hands; if there were disagreements, validity evaluation experts led discussions to refine language. Upon completion of the pictorial representation of the common project ToA, validity evaluation experts drafted the ToA in narrative form for state approval.

Virtual Development Activity
Upon completion of the ToA drafting activities at the SCILLSS Kick-off Meeting, project partners sent the state-specific and common project ToAs and corresponding narratives to the states. States had two weeks to review and provide any additional feedback or suggested revisions to the materials. Project partners refined the state-specific ToAs based on state feedback and subsequently refined the common project ToA to reflect updates made to the state-specific ToAs.

The final state-specific ToAs for Montana, Nebraska, and Wyoming are posted to the SCILLSS website at www.scillsspartners.org. The final SCILLSS Project Theory of Action is provided below.

SCILLSS Project Theory of Action

The SCILLSS project ToA articulates the interrelationships among system components that are necessary to achieve the desired outcomes for both students and educators. By the end of grade 12, all students are critical consumers of information and apply and transfer three-dimensional knowledge and skills in cross-disciplinary ways, demonstrating preparedness for college, the workforce, and civic opportunities. Students demonstrate an appreciation of and engagement with the beauty and wonder of science; possess sufficient knowledge of science and engineering to engage in public discussions on related issues; and continue to learn about science outside of school. The characteristics and priorities of this system also support educators to become effective leaders and to critically implement curriculum, instruction, and assessment (C-I-A) products, processes, and data to support instruction and student learning.

To support these outcomes, students and their families must recognize relationships between school learning, community, and career pathways, and pursue additional science learning experiences and opportunities within and outside of the education system. With educator support, students must take ownership of their science learning by engaging in ongoing progress monitoring using data to track the acquisition of science knowledge and skills.

These student actions will occur if educators cultivate student interest and engagement in content and practices by effectively integrating the three-dimensions in authentic, place-based, and culturally-relevant learning experiences centered on phenomena; appropriately differentiate instruction to ensure all students have the opportunity to access rigorous three-dimensional learning aligned to the Next Generation Science Standards and to build a cohesive understanding of science over time; engage in ongoing, targeted professional development aligned with student outcomes, instructional shifts, and educator needs; and use formative assessment for collecting and evaluating information in real-time.

For educators to achieve these ends, state and local administrators must understand the purpose and intended uses of state assessment scores, and appropriately use assessment data to make

accountability decisions for districts and schools. School leaders and educators must use student performance data from a balanced and coordinated system of assessments to design standards-based curricular resources and to inform regular adjustments to teaching and learning. Administrators must also effectively support the development and implementation of research-based instruction, and personalized, authentic STEM learning experiences. Educators, students, and families must have access to high-quality and research-based strategies, tools, and supports to provide effective learning opportunities for all students. Ongoing and sustained professional development must be provided to support effective research-based practice and allow stakeholders to communicate and collaborate effectively to coordinate the alignment of C-I-A systems.

As a key component of the educational system, the assessment system must be designed to measure student achievement and inform improvements to curriculum and instruction aligned to rigorous and multi-disciplinary college and workforce readiness standards; maximize student engagement and participation through a technology-enhanced, adaptive assessment experience that minimizes testing time and burden; and be accessible, equitable, and culturally appropriate to the widest range of students possible. Educators and other relevant stakeholders must be involved in its development, and have opportunities to communicate and collaborate effectively to coordinate alignment of C-I-A systems.

Assessment results must connect to local curriculum, instruction, and assessment in a coherent, complementary system designed to provide comprehensive coverage of the knowledge, skills, and abilities essential for college and workforce readiness, and include score information that provides timely and actionable student performance data (score reports) that are accessible to a wide range of stakeholders.

A pictorial representation of the SCILLSS Project Theory of Action is provided in Exhibit 1.

Exhibit 1. SCILLSS Project Theory of Action

Statewide Assessment System Setting / Statewide Assessment System Design

The accountability system encourages appropriate, systematic instruction to support inter-disciplinary learning.

Ongoing and sustained professional development supports effective research-based practice and allows stakeholders to communicate and collaborate effectively to coordinate the alignment of curriculum, instruction, and assessment (C-I-A) systems.

The assessment system is designed to measure student achievement and inform improvements to curriculum and instruction aligned to rigorous college and workforce readiness standards.

State science assessments maximize student engagement and participation through a technology-enhanced, adaptive experience that is designed to minimize testing time and burden.

The assessment system reflects Universal Design for Learning principles; assessments are accessible, equitable, and culturally relevant to the widest range of students possible.

Educators and other relevant stakeholders are involved in state assessment development, and communicate and collaborate effectively to coordinate alignment of C-I-A systems.

Educators, students, and families have access to high-quality and research-based strategies, tools, and supports to provide effective learning opportunities for all students.

The assessment system provides timely and actionable student performance data (score reports) that are accessible to a wide range of stakeholders.

State science assessment results connect to local C-I-A in a coherent, complementary system designed to provide comprehensive coverage of the knowledge, skills, and abilities essential for college and workforce readiness.

Data Use

State and local administrators understand the purpose and intended uses of state assessment scores, and appropriately use assessment data to make accountability decisions for districts and schools.
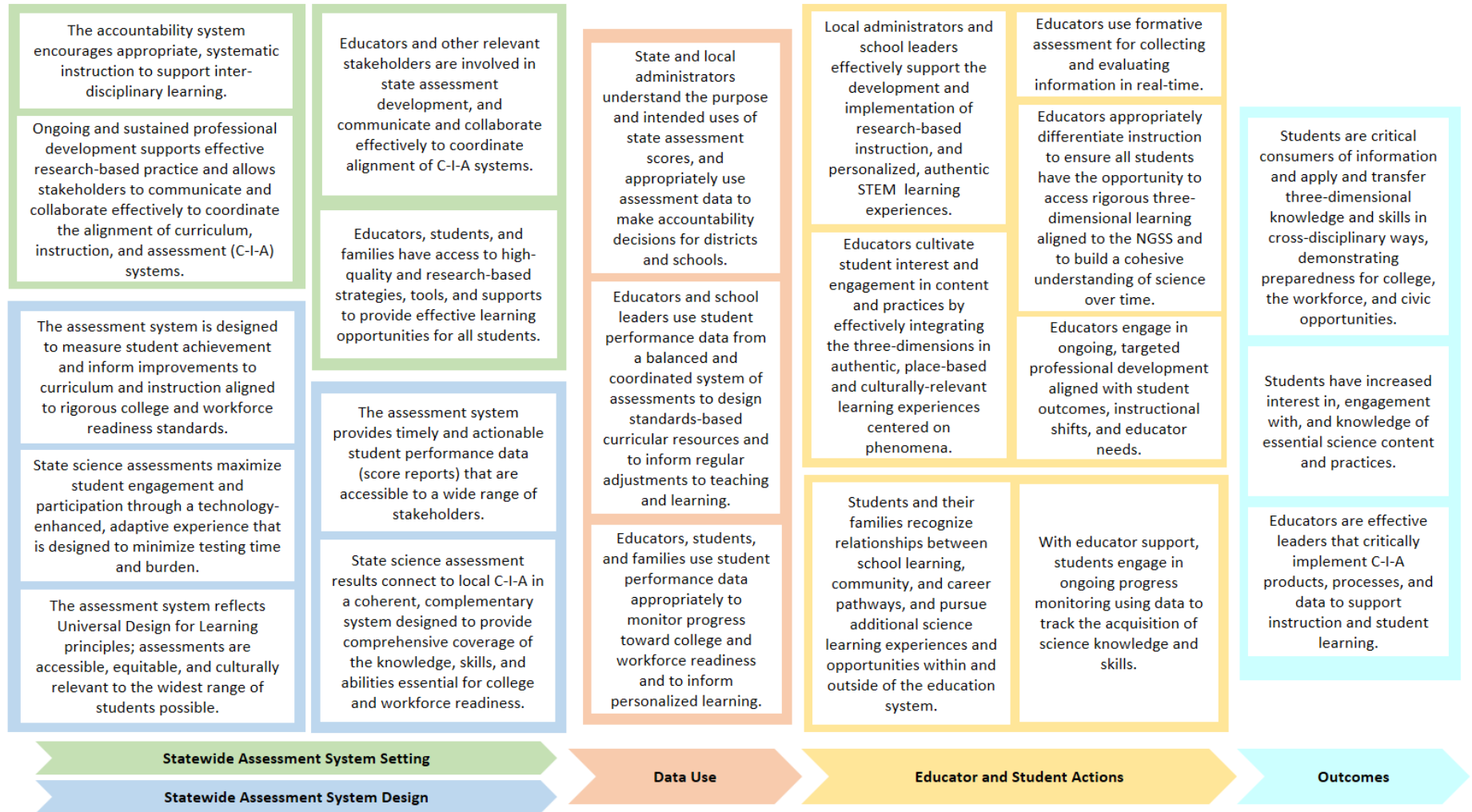
Educators and school leaders use student performance data from a balanced and coordinated system of assessments to design standards-based curricular resources and to inform regular adjustments to teaching and learning.

Educators, students, and families use student performance data appropriately to monitor progress toward college and workforce readiness and to inform personalized learning.

Educator and Student Actions

Local administrators and school leaders effectively support the development and implementation of research-based instruction, and personalized, authentic STEM learning experiences.

Educators cultivate student interest and engagement in content and practices by effectively integrating the three-dimensions in authentic, place-based and culturally-relevant learning experiences centered on phenomena.

Students and their families recognize relationships between school learning, community, and career pathways, and pursue additional science learning experiences and opportunities within and outside of the education system.

Educators use formative assessment for collecting and evaluating information in real-time.

Educators appropriately differentiate instruction to ensure all students have the opportunity to access rigorous three-dimensional learning aligned to the NGSS and to build a cohesive understanding of science over time.

Educators engage in ongoing, targeted professional development aligned with student outcomes, instructional shifts, and educator needs.

With educator support, students engage in ongoing progress monitoring using data to track the acquisition of science knowledge and skills.

Outcomes

Students are critical consumers of information and apply and transfer three-dimensional knowledge and skills in cross-disciplinary ways, demonstrating preparedness for college, the workforce, and civic opportunities.

Students have increased interest in, engagement with, and knowledge of essential science content and practices.

Educators are effective leaders that critically implement C-I-A products, processes, and data to support instruction and student learning.

5

Moving Forward: Sources of Evidence

Once a ToA is in place, the next step is to develop an evidence collection plan and study designs that will yield evidence that could refute or confirm the claims in the state-specific and common project ToA. The purpose of validity evaluation is to test assumptions and produce actionable recommendations; Perie and Forte (2011) call this a "falsification orientation." Thus, the primary focus will be on developing studies that may yield evidence that contradicts claims and assumptions about the assessment system, or that explore alternate explanations for observed phenomena. Ideally, these studies will provide opportunities to address any unique state claims as well.

In total, evidence gathered through these studies should address any critical questions that could be asked of the SCILLSS project by external experts, such as the technical advisory committee (TAC), or by peer reviewers. The findings and evidence gathered via the validity evaluation process ultimately will form the basis for a final validity argument of the SCILLSS project, which is proposed through the ToA, the evidence as provided through project activities and studies, and documentation. Upon approval of the ToAs, project staff will lead states in identifying studies that provide key pieces of evidence of the claims in the ToA relating specifically to the SCILLSS project goals, deliverables, and activities.

References

American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing*. Washington DC: AERA.

Perie, M. & Forte, E. (2011). Developing a validity argument for assessments of students in the margins. In Russell M. (Ed.), *Assessing students in the margins: Challenges, strategies, and techniques* (pp. 335–378). Charlotte, NC: Information Age.