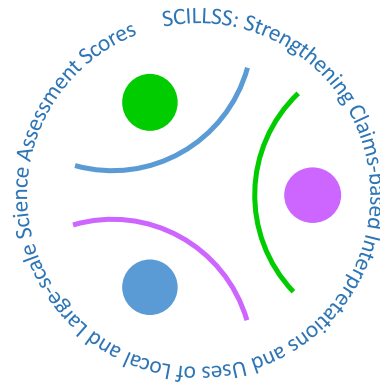


# Strengthening Claims-based Interpretations and Uses of Local and Large-scale Science Assessment Scores (SCILLSS)



## Project Description for the TILSA SCASS

Ellen Forte, edCount, LLC, and Howard Everson, SRI International

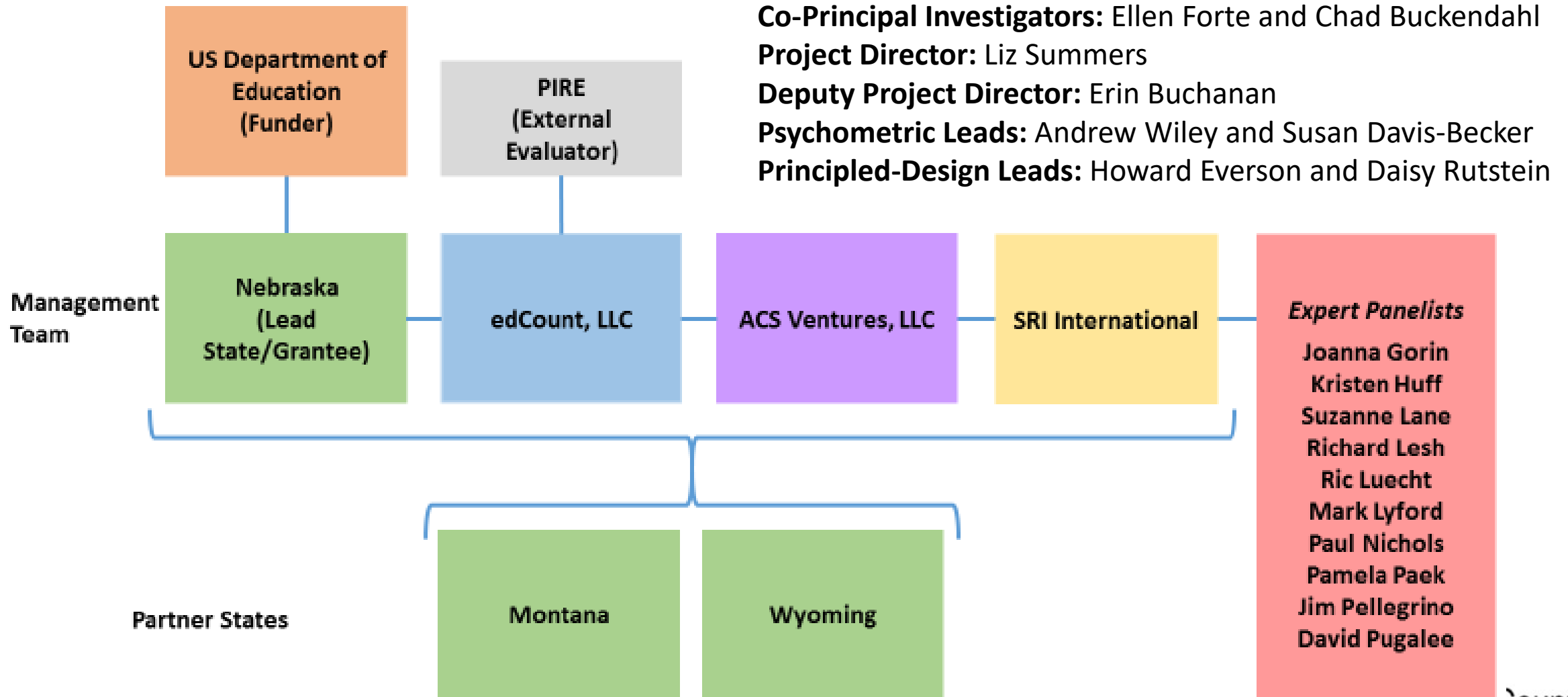
October 24, 2017



# About SCILLSS

- One of two projects funded by the most recent awards under the US Department of Education's Enhanced Assessment Instruments Grant Program (EAG), announced in December, 2016
- Collaborative partnership including three states, four organizations, and 10 expert panel members
- Four year timeline (April 2017 – December 2020)

# SCILLSS Partner States, Organizations, and Staff



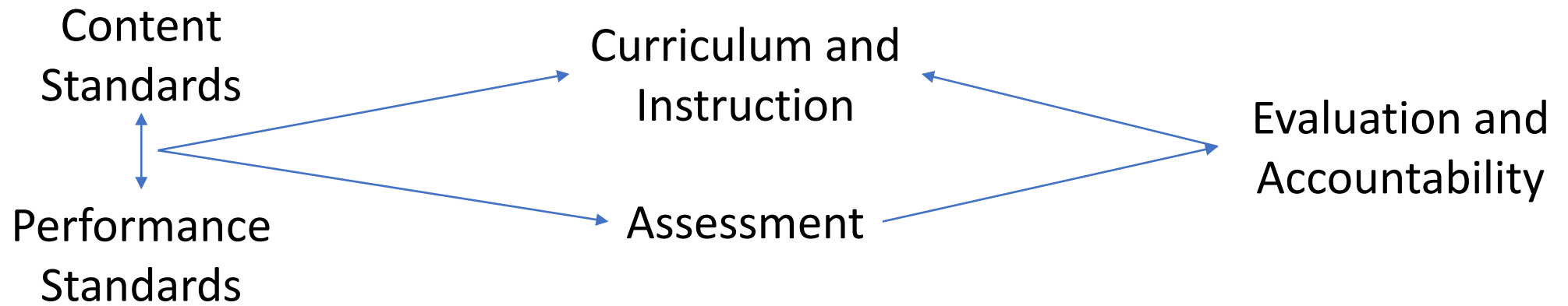
# SCILLSS Project Goals



- Create an assessment design model that **enhances alignment with standards by eliciting common construct definitions** that drive curriculum, instruction, and assessment
- Strengthen a shared knowledge base among stakeholders for using **principled-design approaches** to create and evaluate quality science assessments that **generate meaningful and useful scores**
- Establish a means for state and local educators to **connect statewide assessment results with local assessments and instruction** in a coherent, standards-based system

# Score Interpretations for Standards-Based Assessments

- Systemic reform model



- Standards are the domain
- Curricula and assessments are interpretations of the domain
- If the assessment interpretation is poor - or significantly different from the curricular interpretation - then the scores don't reflect the domain and the model falls apart

# 5 Components of Performance Standards

More sophisticated performance



Less sophisticated performance

Levels

## **Name of most sophisticated level**

- Description of performance in this level
- Samples of student work

## **Name of next most sophisticated level**

- Description of performance in this level
- Samples of student work

## **Name of next least sophisticated level**

- Description of performance in this level
- Samples of student work

## **Name of least sophisticated level**

- Description of performance in this level
- Samples of student work

Cut scores

## Then scores can be used...

### If scores reflect what students...

know and can do

have learned this year/in this course

to track student progress toward...

for school accountability decisions and program evaluation

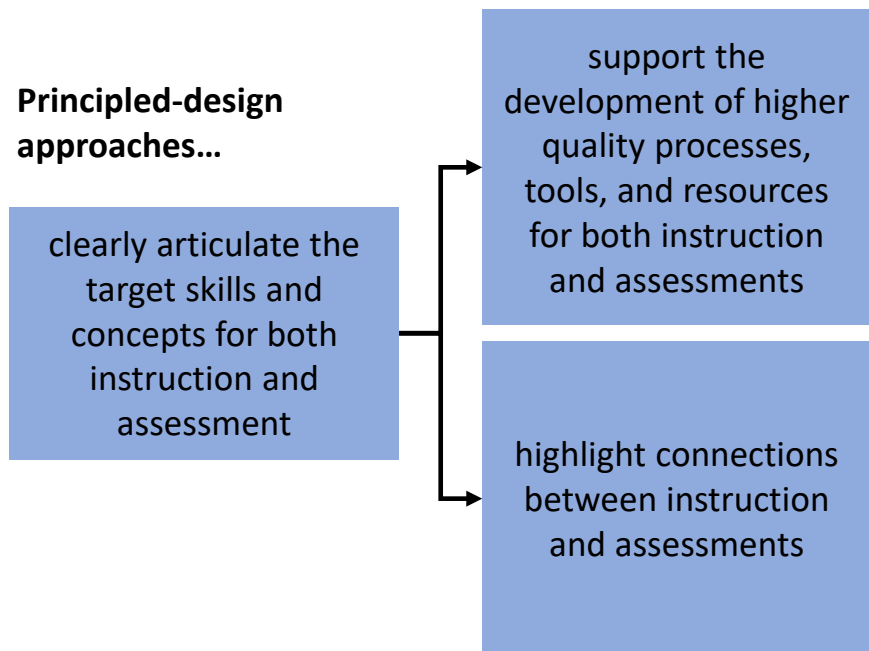
to build and deliver instruction aligned with academic expectations

### IF

- **Constructs are well-defined**
- **Construct definitions are shared across the system**
- **The system is well-designed**
- **The system is well-implemented**

- Constructs are well-defined
- Construct definitions are shared across the system
- The system is well-designed
- The system is well-implemented

**Principled-design approaches...**



**and this enhances...**

- interpretability of assessment scores
- appropriate use of assessment scores
- clarity of claims-based instruction and classroom assessment

**which requires...**

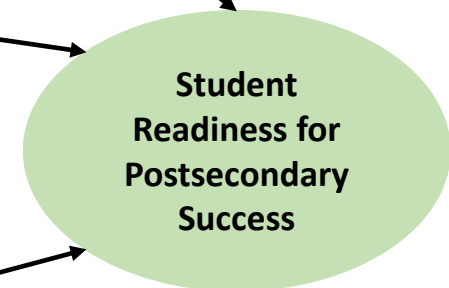
- collaborative engagement of state and local educators and measurement experts

**which lead to...**

- efficiencies in test selection and use
- coherence in relationships among standards, instruction, and assessment
- coherence in relationships among components in assessment systems

**and improves...**

- communication and connections among educators and other stakeholders

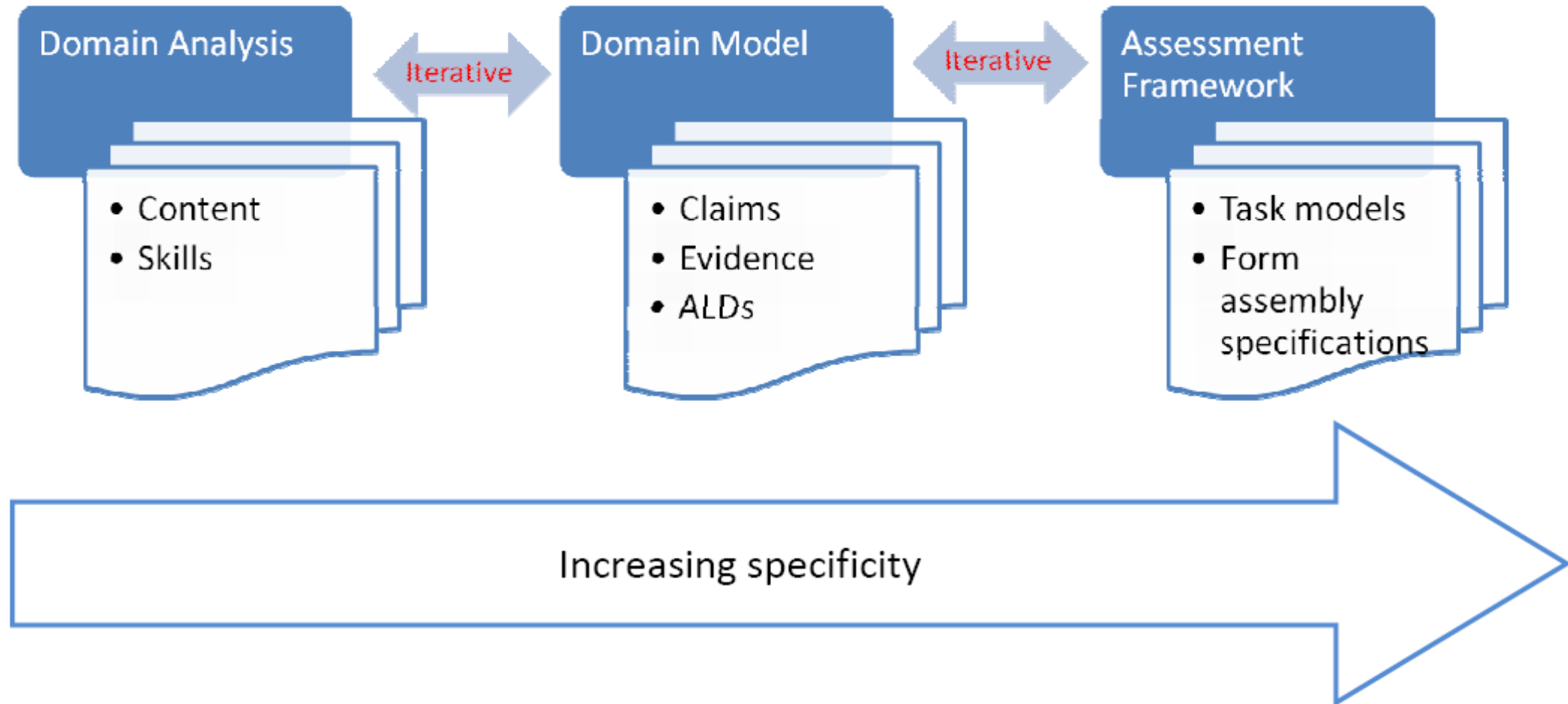




# Benefits of a Principled-Design Approach

- Principled articulation and alignment of design components
- Articulation of a clear assessment argument
- Reuse of extensive libraries of design templates
- For accountability
  - Clear warrants for claims about what students know and can do
  - Build accessibility into design of tasks (not retrofitted into tasks)
  - Cost v. scale

# Three Critical Evidence-Centered Design Phases



Adapted from Huff, Steinberg, & Matts, 2010

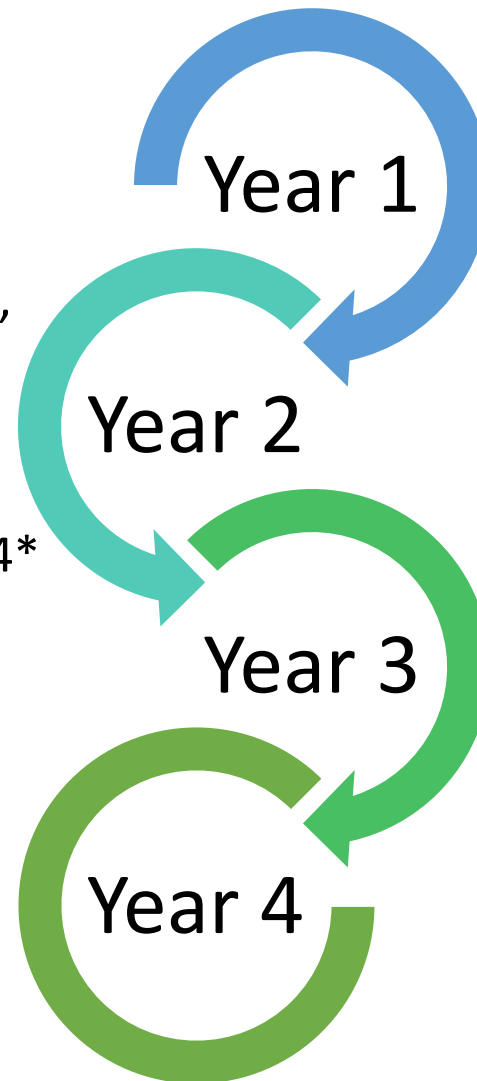
# Project Deliverables

## 2 - Large-scale assessment resources

- For each prioritized claim:
  - PLDs
  - Measurement targets, task models, and design patterns
  - Sample items for each design pattern
- Assessment literacy modules 2, 3\* and 4\*

## 4 - Reporting and Dissemination

- Database of student artifacts corresponding to the performance levels
- Post-project survey
- Post-project action plans for each state
- Final project report



## 1 - Project Foundations

- SCILLSS website
- Local and state needs assessment tools
- Theory of Action (ToA) template
- Common project ToA
- Three state-specific ToAs
- Assessment literacy module #1
- Three prioritized science claims

## 3 - Classroom-based assessment resources:

- Six task models
- Six tasks
- Six sets of student artifacts

# Theory of Action 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)?

# Theory of Action Progress

- Completed
  - Drafted initial versions of state-specific and project ToAs
  - States reviewed state-specific and project ToAs and provided feedback to inform revisions
- Next steps
  - Revise ToAs based on state feedback and ensure project ToA reflects common priorities of participating states
  - Document ToA development process and prepare ToAs for dissemination via project website



# Needs Assessment Design: 3 Steps

Identify

Identify uses and purposes of assessment data in your school/district/state

Evaluate

Gather and evaluate evidence for each assessment related to four overarching validity questions

Synthesize

Synthesize results from steps 1 and 2 to determine an appropriate path forward

# Four Overarching Validity Questions

1. To what extent does the assessment as designed and implemented capture the knowledge and skills defined in the target domain?
2. To what extent does the assessment as implemented yield scores that are comparable across students, sites, time, forms?
3. To what extent are students able to demonstrate what they know and can do in relation to the target knowledge and skills on the test in a manner that can be recognized and accurately scored?
4. To what extent does the test yield information that can be and is used appropriately within a system to achieve specific goals?

**Construct Coherence**

**Comparability**

**Accessibility and  
Fairness**

**Consequences**

# Needs Assessment Progress

- Completed
  - Collaborated to develop initial versions of the local needs assessment
  - Designed the first assessment literacy module to complement and expand upon topics introduced in the needs assessment
- Next steps
  - State leads and TAC members review the needs assessment and provide feedback to inform revisions
  - Pilot the needs assessment with school districts in each participating state





# Assessment Literacy Module One

- 1) Discuss the purposes and uses of assessment scores
- 2) Present validity as the key principle of assessment quality
- 3) Describe the phases of the assessment life cycle:
  - Design and development
  - Administration
  - Scoring
  - Analysis
  - Reporting
- 4) Discuss four questions that cover the breadth of the validity issues responsible test users must consider:
  - Construct coherence
  - Comparability
  - Accessibility and fairness
  - Consequences

# Assessment Literacy Module Progress

- Completed
  - Researched and confirmed module delivery platform
  - Drafted presentation outline and template for module 1
  - Drafted module 1 content and interactive features
  - Designed module 1 to complement and expand upon topics introduced in the needs assessment
- Next steps
  - Complete internal review
  - TAC review of the module to provide feedback for revisions
  - Finalize, publish, and pilot module 1 in tandem with local needs assessment



# District and Educator Involvement



Implement Local Needs Assessment (Self-evaluation Tool)



Educator Review of Large-Scale Assessment Materials: PLDs, Measurement Targets, Task Models, and Items



Pilot Study of Classroom-based Evidence and Tools