Considering the Design and Role of Interim Assessments in an NGSS System of Assessments

Nathan Dadey & Sara Cooper

The Center for Assessment & The Nebraska Department of Education

NCME Classroom Assessment Conference

Lawrence, KS

October 8, 2018

Center for Assessment

Overview

1. A quick re-introduction to the Next Generation Science Standards
2. Thinking about state led design of district assessments.
3. Nebraska's development efforts on district assessments.

Center for Assessment

1 | Reviewing the NGSS

The Next Generation Science Standards translates **three dimensions** of high quality science **instruction** presented in the Framework for K-12 Science Education

Into

The "**performances** students must be able to do at the **conclusion of instruction**".

I.e., **Performance Expectations** (PEs).

[1]See The Next Generation Science Standards: Introduction, p. 3

Motion and Stability: Forces and Interactions

Students who demonstrate understanding can:

Performance Expectations

Support an argument that the gravitational force exerted by Earth on objects is directed down. 5-PS2-1

▸ Clarification Statement and Assessment Boundary

Science and Engineering Practices

Engaging in Argument from Evidence

Engaging in argument from evidence in 3–5 builds on K–2 experiences and progresses to critiquing the scientific explanations or solutions proposed by peers by citing relevant evidence about the natural and designed world(s).

- Support an argument with evidence, data, or a model. (5-PS2-1)

Disciplinary Core Ideas

PS2.B: Types of Interactions

- The gravitational force of Earth acting on an object near Earth's surface pulls that object toward the planet's center. (5-PS2-1)

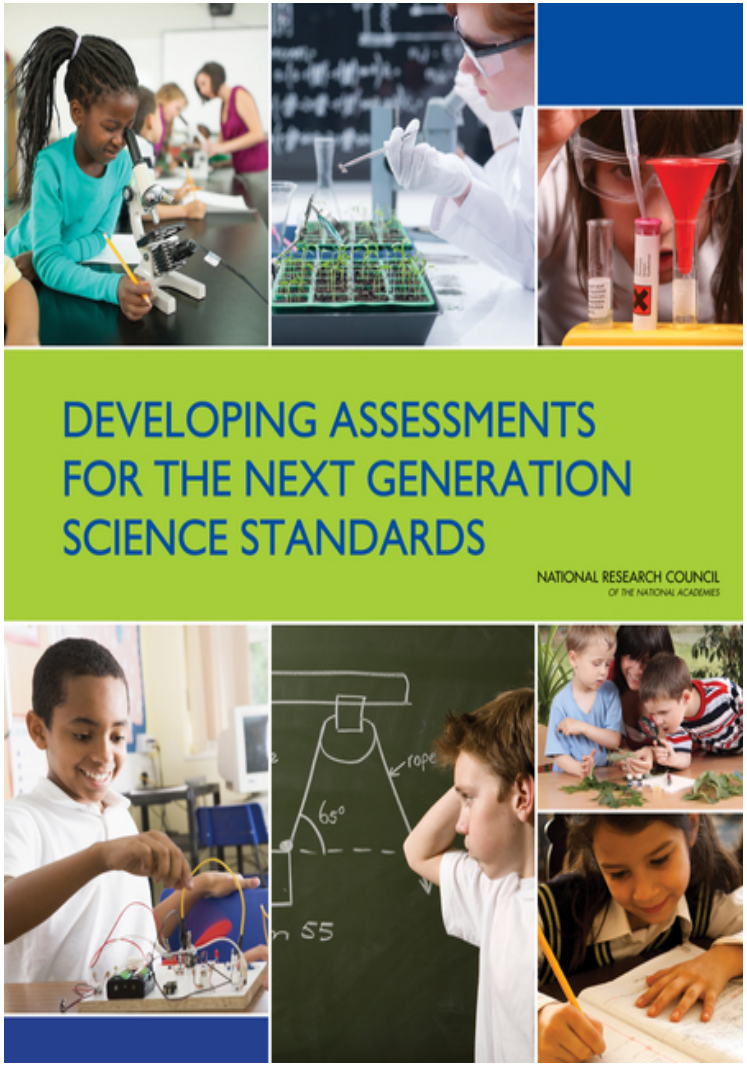
Crosscutting Concepts

Cause and Effect

- Cause and effect relationships are routinely identified and used to explain change. (5-PS2-1)

https://ngss.nsta.org/DisplayStandard.aspx?view=pe&id=110

Systems & the NGSS

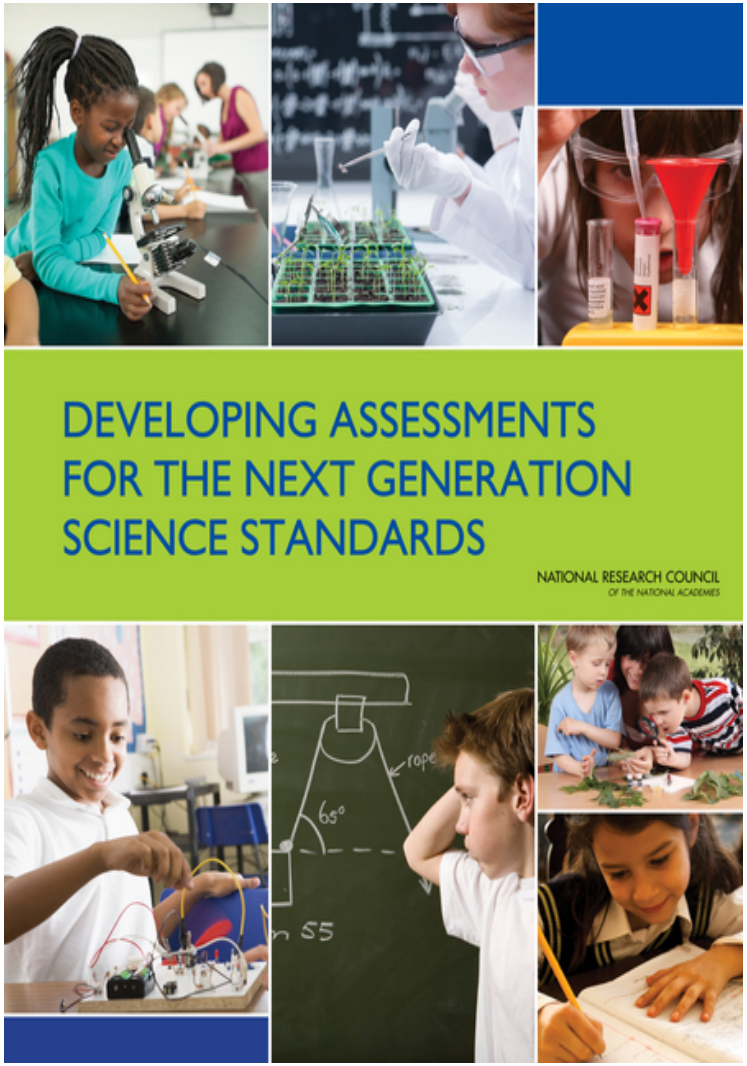


National Research Council, 2014

"It will not be feasible to assess all of the performance expectations for a given grade level during a single assessment occasion. Students will need multiple—and varied—assessment opportunities."

Opportunities embedded in a **system of assessments**.

Systems & the NGSS



National Research Council, 2014

A **system of assessments** will be needed to measure the NGSS performance expectations and provide students, teachers, administrators, policy makers, and the public with the information each needs about student learning.

2 | Considering District-Level Design

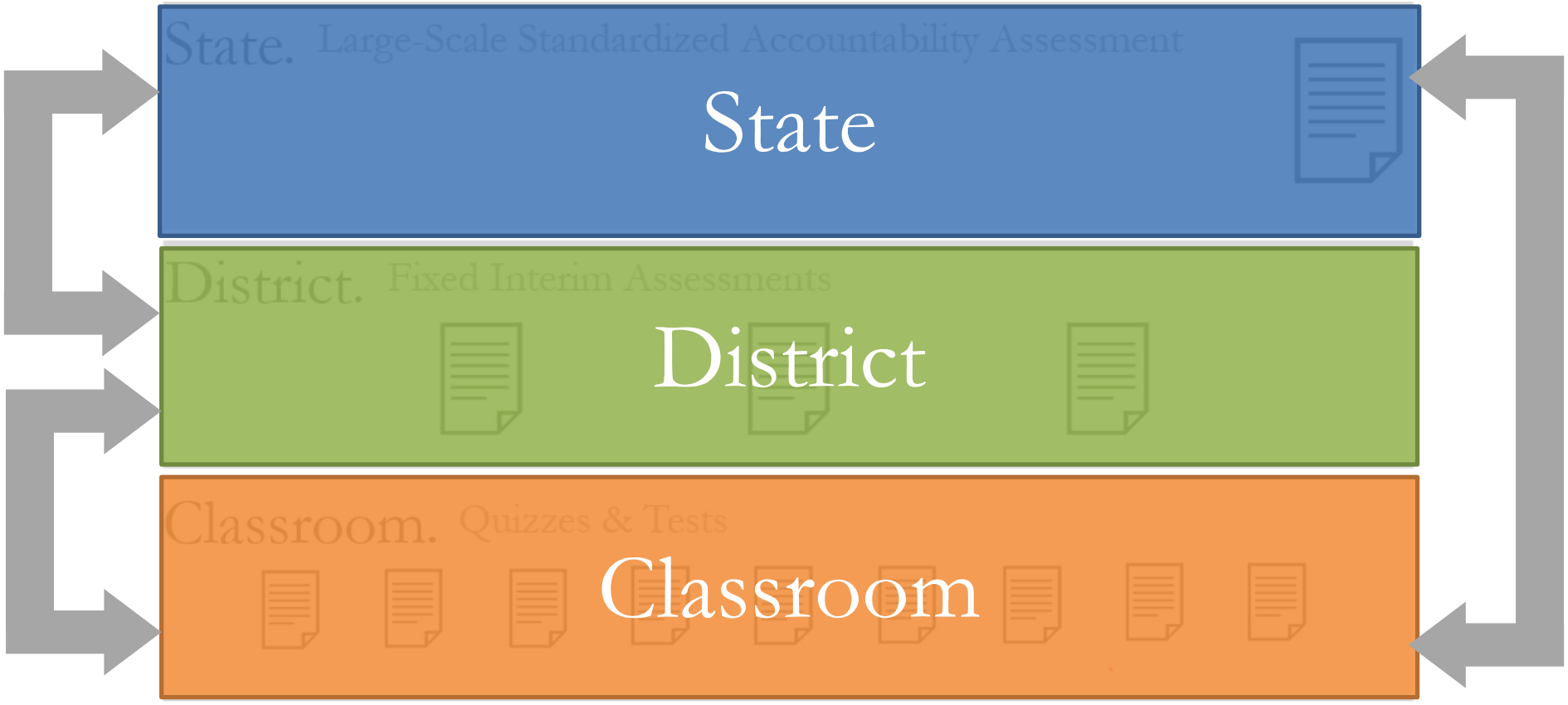
Context (Not NGSS Specific)

Some states have:

- increasingly taken interest in interim assessments[1], perhaps in response to their continued wide-spread use, as well as federal attention.
- developed programs aimed at providing interim assessments or improving practice around interim assessment (i.e., a top down approach).
  - Generally targeted to the district level, although schools could leverage many of the same resources.

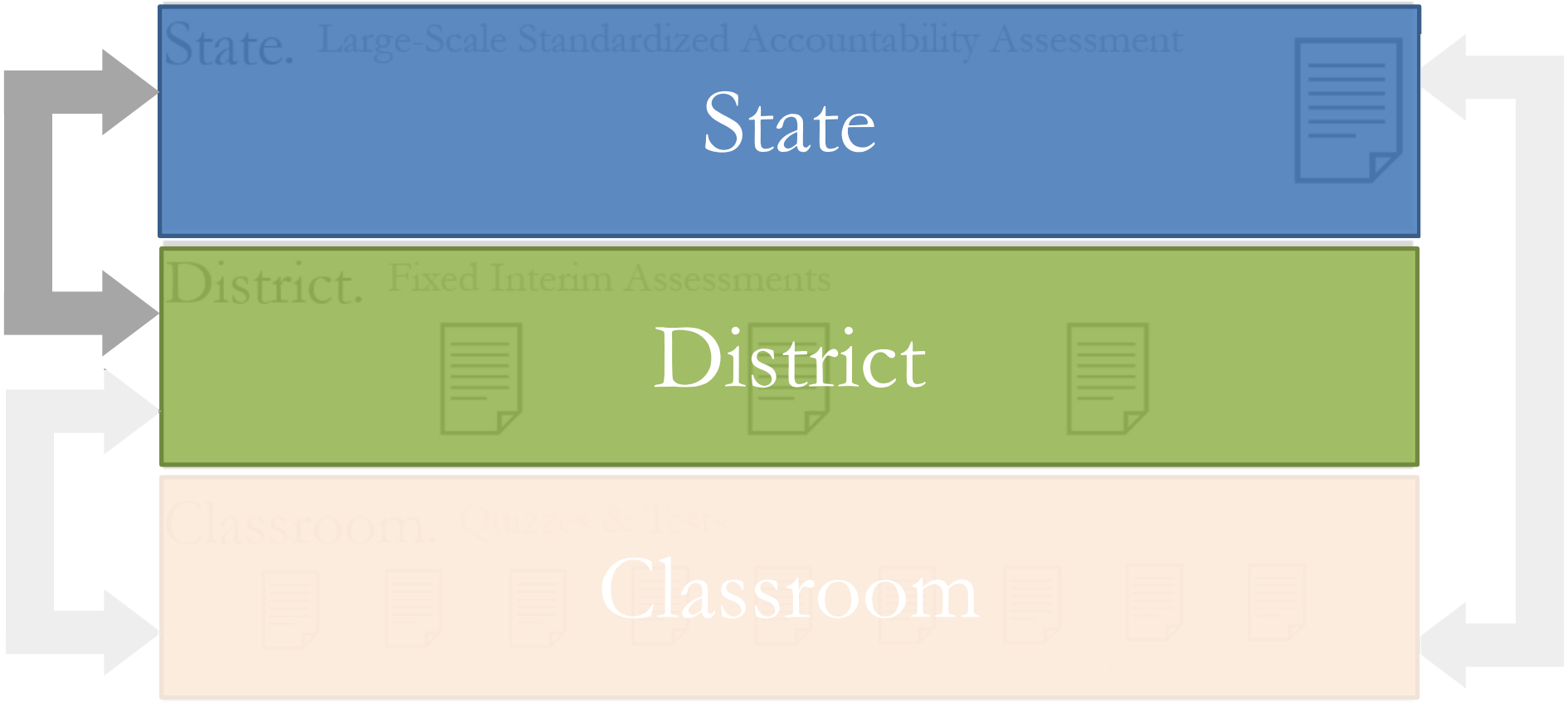
[1]As defined by Perie et al. (2009): "assessments administered during instruction to evaluate students' knowledge and skills relative to a specific set of academic goals in order to inform policymaker or educator decisions at the classroom, school, or district level. The specific interim assessment 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" (2009, p. 6, emphasis added).

Considering Levels of a System



Note: Not all levels are considered here, others could be more relevant, depending on context.

Considering Levels of a System



The focus of this presentation is on the way in which states can support district practice.

Note: Not all levels are considered here, others could be more relevant, depending on context.

Systems Perspective

- Considering how these efforts function from a system of assessment perspective leads us to ask questions like:
  - What is the relationship between the state- and district-level assessments?
  - How are the assessments complementing one another in terms of coherence, comprehensiveness and continuousness (plus utility & efficiency)?

Additional Design Considerations

Addressing these types of questions leads to two additional, overlapping criteria that could be helpful:

- **Learning Targets**: what are the depth and breath of the interim assessment learning targets, in relation to the state-level assessments (and vice-versa)? (Gong, 2010)
- **Degree of Modularity**: if there are multiple assessments:
  - At what grain size is the content targeted to?
  - How flexible is the timing of administration?

[1]In this presentation, the terms interim assessments and district-level assessment are interchangeable.

Examining Current Practice

- As a start, reviewed 20 of 50 state department of education assessment websites (Alabama to Maryland)
- To see how these issues of design have been addressed and what trends exist.
  - Degree of modularity was easily captured, so we start there
  - Learning targets is something we are looking forward to investigating further

Degree of Modularity

Fixed Designs

A single assessment that measures the entire domain

Block Designs

Multiple assessments, each measuring a chunk of the domain

Modular Designs

Multiple assessments, each measuring a very small chunk of the domain

Fixed

Modular

Fixed Designs (8 states)

Fixed Designs

A single assessment that measures the entire domain

- Broadly measure the domain (sometimes in ways that mirror the state-level summative accountability assessment)
- Administered as needed or within specified administration windows
- Examples:
  - Smarter Balanced Interim Comprehensive Assessments
  - KS Interim Predictive Tests
  - LA LEAP 360 Diagnostic & Interim

Fixed Designs (8 states)

Name/Type of Assessment	# States	States
Smarter Balanced Interim Comprehensive Assessments	5	CA, CT, DE, HI, ID
Early Literacy Assessments	1	IN, CO*
Interim Predictive Tests	1	KS
LEAP 360 diagnostic & interim	1	LA

Block Designs (6 States)

Block Designs

Multiple assessments, each measuring a chunk of the domain

- Measures a predefined set of the domain (e.g., content domain, unit grouping)
- Administered on demand, sometimes within pre-specified windows
- Examples:
  - Delaware's end of unit assessments
  - Louisiana's EAG ELA assessments
  - Wyoming's modular interim assessments*
  - Smarter Balanced Interim Assessment Blocks

*Not included in counts.

Block Designs (6 states)

Name/Type of Assessment	# States	States
Smarter Balanced Interim Assessment Blocks	5	CA, CT, DE, HI, ID
Unit Based Assessments	2	DE, LA

Modular Designs (5 States)

Modular Designs

Multiple assessments, each measuring a very small chunk of the domain

- Measures a fine grained part of the domain
- Administered on demand, sometimes within pre-specified window, and sometimes with recommended groupings
- Examples:
  - Wyoming's modular assessments
  - Kentucky's through course tasks

*Not included in counts.

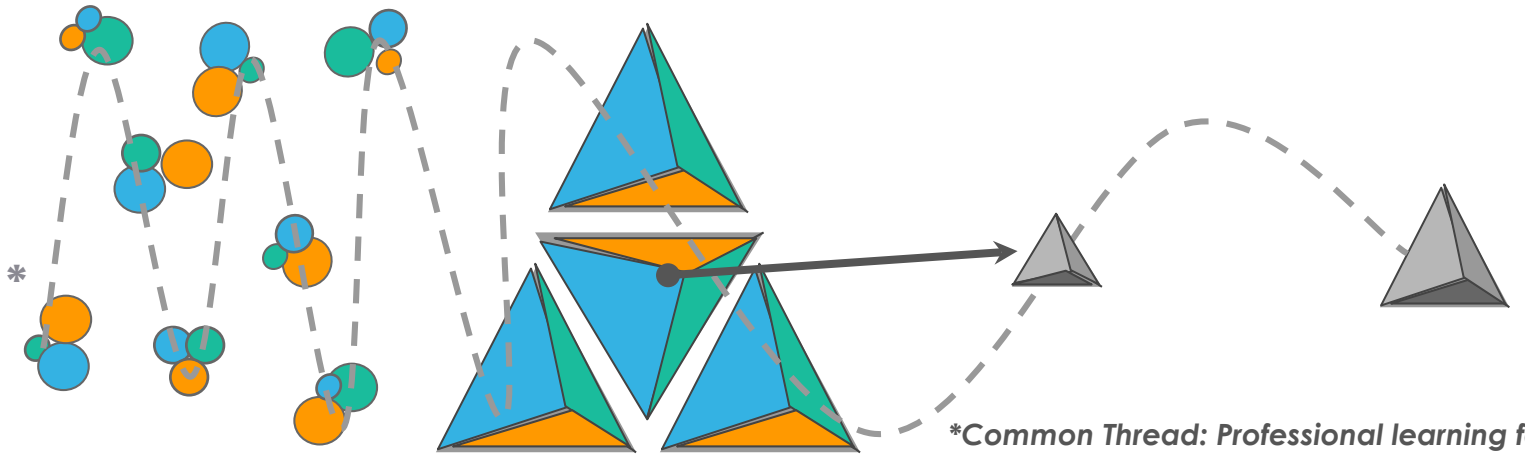
Modular Designs (6 states)

Name/Type of Assessment	# States	States
Task Banks	3	FL, LA, CT, GA
“Standards” Based Assessments	2	KS, GA

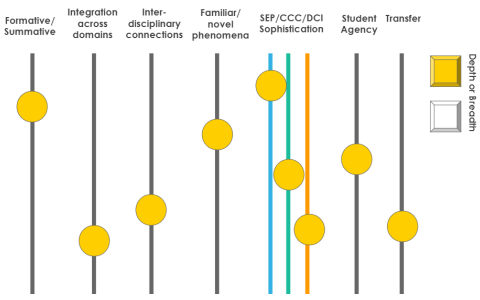
3 | Nebraska's Efforts

Nebraska's Balanced Science Assessment System Model

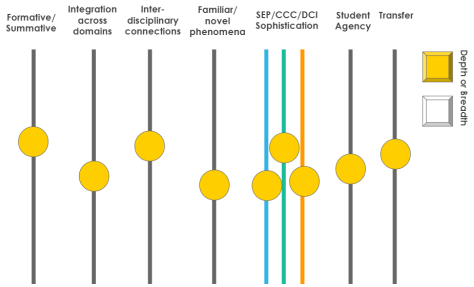
Phenomena and Problem-focused	Engage diverse sense-making	Require reasoning with evidence	Grade appropriate 3D targets	Demonstrate science under-standing by doing science



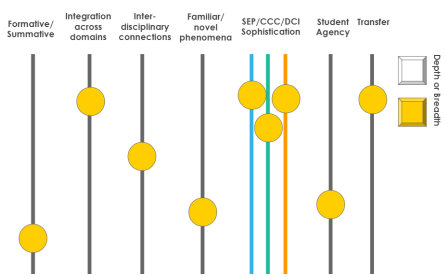
Common Thread: Professional learning for educators



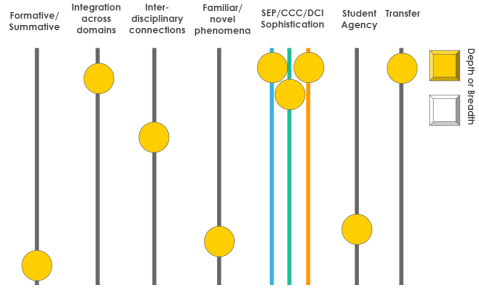
Formative
(K-12)



Interim
(K-12)

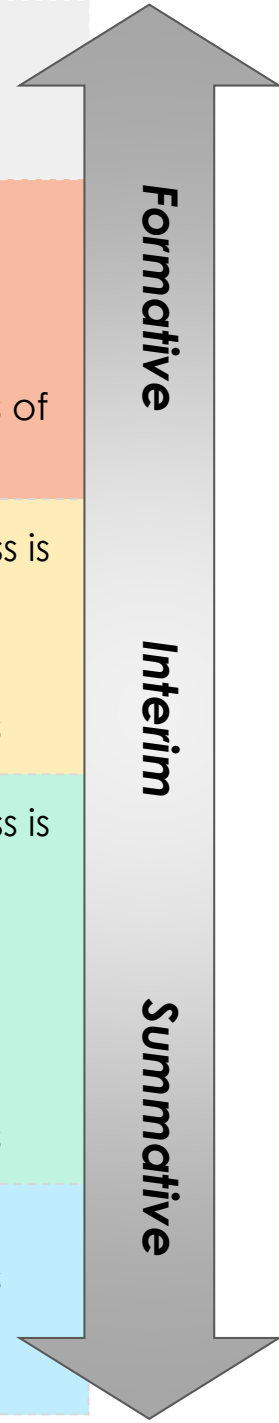


Monitoring
(3,4,6,7,9,10)

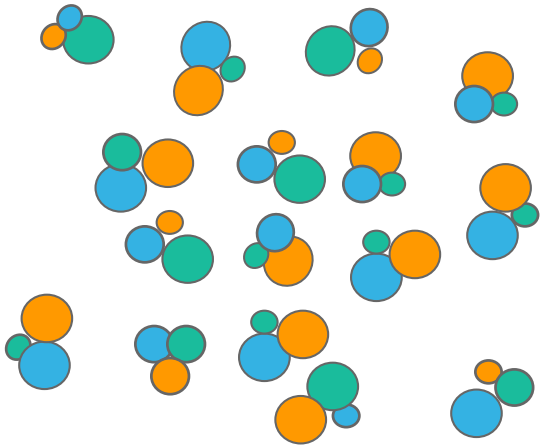


Summative
(5, 8, 11*)

	Purpose	*Information*	*Frequency and Method*	*Decision Level(s)*	*Use/Actions*
Classroom Embedded Assessments	quickly inform instruction	specific, immediate, actionable	daily, ongoing instructional strategies	student/ classroom	what comes next for students effectiveness of instruction
Interim Task Library	benchmark and monitor progress	multiple data points over time	periodic common assessments	grade level/ school/ district	what progress is being made program effectiveness
Complementary Statewide Summative	benchmark and monitor progress evaluate learning	multiple data points over time	annual standardized assessments	grade level/ school/ district/ state	what progress is being made student proficiency program effectiveness
Statewide Summative	evaluate learning	cumulative snapshots	triennial standardized assessments	school/ district/ state	program effectiveness



Design Challenges & Solutions



Classroom

- Diverse local curricula
- Degree of educator capacity for 3D teaching and assessment

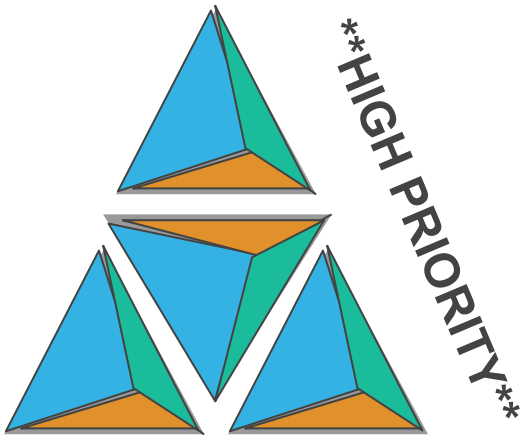
Replacement Units

- Model short instructional sequence
- embedded formative assessment probes
- end of unit task

Professional Learning Modules

- Online platform
- Co-developed with the Science Cadre & SCILLSS

Design Challenges & Solutions



Interim

- Development of tasks
  - student work samples
- Need to support a range of uses
- Signal to the classroom level and to summative
- Valid and reliable scoring

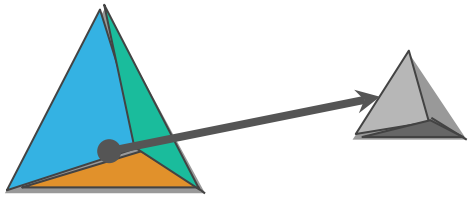
Interim Assessment Development Workshops

- Co-facilitated by NDE and ESUs or individual districts
- Collaborate across states
- Engage ***all*** stakeholder groups

Professional Learning Modules

- Online platform
- Co-developed with the Science Cadre

Design Challenges & Solutions



Monitoring

- Development of tasks
- Platform/test engine
- Integrate across content
- Achievement? and growth metrics?
- Psychometric model for rolling scores
- Test security
- Multiple high school course models

??????????????

- Collaborate across states
- Engage ***all*** stakeholder groups
- Domain level 'choice' for districts at HS level

Professional Learning Modules???

Design Challenges & Solutions

Summative

- Phenomena/problem driven performance tasks
- NE interpretation of standards
- Interdisciplinary integration
- Choice/student agency
- Transparent scope
- Signal to the system; not student
- Rigorous alignment
- Opportunity to learn
- Educator capacity for development

Working Solutions

- Collaboration
- 1 overall student score
- Flexible blueprint
- Communication plan

Professional Learning

- Embedded in assessment development workshop

Challenges

- Fire-walling purposes and uses
- Peer Review
- Funding and Longevity
- Technical Platform