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
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.
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

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

---

1See [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)

“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.
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
Some states have:

- increasingly taken interest in interim assessments\(^1\), perhaps in response to their continued widespread 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.
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.
• 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)?
Addressing these types of questions leads to two additional, overlapping criteria that could be helpful:

• **Learning Targets**: what are the depth and breadth 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?

1In 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
<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Designs</td>
<td>A single assessment that measures the entire domain</td>
</tr>
<tr>
<td>Block Designs</td>
<td>Multiple assessments, each measuring a chunk of the domain</td>
</tr>
<tr>
<td>Modular Designs</td>
<td>Multiple assessments, each measuring a very small chunk of the domain</td>
</tr>
</tbody>
</table>
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)

<table>
<thead>
<tr>
<th>Name/Type of Assessment</th>
<th># States</th>
<th>States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smarter Balanced Interim Comprehensive Assessments</td>
<td>5</td>
<td>CA, CT, DE, HI, ID</td>
</tr>
<tr>
<td>Early Literacy Assessments</td>
<td>1</td>
<td>IN, CO*</td>
</tr>
<tr>
<td>Interim Predictive Tests</td>
<td>1</td>
<td>KS</td>
</tr>
<tr>
<td>LEAP 360 diagnostic &amp; interim</td>
<td>1</td>
<td>LA</td>
</tr>
</tbody>
</table>
Block Designs (6 States)

- 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)

<table>
<thead>
<tr>
<th>Name/Type of Assessment</th>
<th># States</th>
<th>States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smarter Balanced Interim Assessment Blocks</td>
<td>5</td>
<td>CA, CT, DE, HI, ID</td>
</tr>
<tr>
<td>Unit Based Assessments</td>
<td>2</td>
<td>DE, LA</td>
</tr>
</tbody>
</table>
Modular Designs (5 States)

- 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.*
<table>
<thead>
<tr>
<th>Name/Type of Assessment</th>
<th># States</th>
<th>States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Banks</td>
<td>3</td>
<td>FL, LA, CT, GA</td>
</tr>
<tr>
<td>“Standards” Based Assessments</td>
<td>2</td>
<td>KS, GA</td>
</tr>
</tbody>
</table>
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 understanding 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*)
<table>
<thead>
<tr>
<th></th>
<th>Purpose</th>
<th>Information</th>
<th>Frequency and Method</th>
<th>Decision Level(s)</th>
<th>Use/Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classroom Embedded</strong></td>
<td>quickly inform instruction</td>
<td>specific, immediate, actionable</td>
<td>daily, ongoing instructional strategies</td>
<td>student/classroom</td>
<td>what comes next for students</td>
</tr>
<tr>
<td><strong>Assessments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>effectiveness of instruction</td>
</tr>
<tr>
<td><strong>Interim Task</strong></td>
<td>benchmark and monitor progress</td>
<td>multiple data points over time</td>
<td>periodic common assessments</td>
<td>grade level/school/district</td>
<td>what progress is being made</td>
</tr>
<tr>
<td><strong>Library</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>program effectiveness</td>
</tr>
<tr>
<td><strong>Complementary</strong></td>
<td>benchmark and monitor progress</td>
<td>multiple data points over time</td>
<td>annual standardized assessments</td>
<td>grade level/school/district/state</td>
<td>what progress is being made</td>
</tr>
<tr>
<td><strong>Statewide Summative</strong></td>
<td>evaluate learning</td>
<td>cumulative snapshots</td>
<td>triennial standardized assessments</td>
<td>school/district/state</td>
<td>student proficiency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>program effectiveness</td>
</tr>
<tr>
<td><strong>Statewide Summative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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
Interim Assessment

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

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

Professional Learning Modules
- Online platform
- Co-developed with the Science Cadre
Design Challenges & Solutions

- 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

Monitoring

- 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