Assessing Student Learning of the Next Generation Science Standards

Scott Marion
Center for Assessment

#RILS2017 Reidy Interactive Lecture Series
Portsmouth, NH
September 28-29, 2017
Named for a famous Kentucky educational leader, Ed Reidy, RILS brings together **participants** with a range of expertise to wrestle with difficult challenges in search of practical solutions or promising approaches. Participants are encouraged to participate!
Who’s Here? Lots of RILS veterans & newcomers!

• **States**
  – CT, DE, KY, LA, ME, MA, MI, MT, NE, NH, NM, OR, RI, WY

• **Districts**
  – Goffstown, NH; Gwinnett Co., GA; New York, NY; SAU 39, NH

• **Teachers and Principal**— Newmarket and Rye, NH

• **Assessment Companies**
  – AIR, DRC, ETS, GA Center for Assessment, KU-CETE, KU-DLM, Measured Progress, NWEA, Pearson, WestEd

• **Consulting Firms/TA Providers/Advocates**
  – Achieve, Center for Assessment, EdCount, HumRRO, Maine Math/Science Alliance, NSTA, Next Gen Consulting

• **Universities/Research Institutions**
  – Boston College, NORC, SRI, University of Colorado, University of Kentucky, University or Oregon

• **Center for Assessment Board of Directors**
  – Mark Musick, Laurie Wise, Henry Braun, Peter McWalters

• **Retired!**
It Takes a Village!

• Nathan Dadey
• LauraLee McGuane & Sandi Chaplin
• Center associates and staff
• Presenters, panelists, facilitators, and many participants
• The Center’s Board of Trustees
How did we get here?

Late 1980s – 1990s

2010 – 2015
Much of the early NGSS work focuses on **classroom** & district assessment.

Much of the recent work on NGSS assessments focuses primarily on **district** & state assessment.
Now we’re here...

- How many of you designing new science assessments feel like Sisyphus?
- Solving one problem often opens up new challenges...
- Nathan and I felt similarly challenged when designing this program—every topic could be the subject of a week-long conference!
New insights from the NGSS...

- Bill and TJ will help us dive into the NGSS...
- This will re-ground us in the three-dimensional structure of the standards using a phenomenon-based unit
- This 3D structure opens up instructional possibilities, but also creates measurement challenges...
**Claims, Constraints and Requirements**

- All assessment design is a case of optimization under constraints.
- This richness of the NGSS makes these constraints (e.g. time, money) seem especially tight.
- First, Brian, April, Kevin, and Kathleen will discuss how states have developed their major claims and blueprints that the assessment must support.
- Next, Nathan, Jon, Kevin, and TJ will lead us through an activity to think about subclaims and what we want to report to our stakeholders.
Items and Tasks

- We’ll get to do a little window shopping after lunch
- Several participants have graciously agreed to share their items and tasks for us to examine

*Not exactly true, but a good indicator*
Alignment

• Alignment is a “deal-breaker” criterion for U.S. Department of Education’s peer review process

• The NGSS are big! Trying to represent the NGSS on a summative assessment in order to meet typical alignment criteria appears to be a considerable challenge

• Some question whether current alignment methods are capable of allowing fair evaluations of the alignment of NGSS assessment

Aneesha, Sara, Chris, TJ, Jan, and Sean will help us think about these challenges in conceptual, methodological, and practical ways
• Many assessment leaders have been focused on initial development tasks (e.g., item-development, blueprints, alignment), but there are some critical measurement issues lurking just below the surface.

• Joseph, Leslie, Jon and Kathleen will help us consider a few key measurement issues related to NGSS assessments.
The National Research Council’s committee that produced *Developing Assessments for the Next Generation Science Standards* strongly recommended developing a balanced systems of assessments to support the multiple uses a purposes.

Bill and I will discuss the committee’s (Kathleen was also a member) recommendations and discuss opportunities for assessment systems with April and Sean.
Validity Considerations

• It’s all about validity!

• The entire conference is focused on the degree to which our items and tests support inferences about what students know and can do in terms of the NGSS

• While we do not have a separate validity session, it will be one of focal topics in our last session of the conference

• As a preview of the complexity, we offer some insights from work by DeBerger (along with our colleagues Bill and Chris)
Claim: What Students Know and Can Do

Student Assessment Performance

Program Quality

Inference

Assessment Argument

From: DeBarger, Penuel, Harris & Kennedy (2015).
From: DeBarger, Penuel, Harris & Kennedy (2015).
Validity

A lot to think about, but try to keep in mind:

• What is the evidence and logic to support the claims about what the assessment results can tell us about what students know and can do in terms learning 3D science standards?
Breakfast Tables Friday Morning @7:30

1. Measurement models and scale maintenance
2. ESSA’s Innovative Assessment Pilot: The Potential for Science Assessment
3. The state of standard setting
4. Measuring writing in response to text – does this require reading comprehension, essay writing, or a different construct of “analysis”
5. Evaluating accountability systems for identification and continuous improvement
Join us for reception this evening at 5:30
• Respond to survey at the end of the day today to help us dynamically create groups for our last session of the conference
Media

- Live tweeting #RILS2017 @NCIEA1
- Photos
- Video
join THE CENTER FOR ASSESSMENT

for our 20th Anniversary Celebration and Edward F. Reidy, Jr. Interactive Lecture Series

September 26-28, 2018

A Look Back and a Look Ahead at Educational Assessment and Accountability