

# 2017 Program in the Edward F. Reidy, Jr. Interactive Lecture Series

---

## Assessing Student Learning of the Next Generation Science Standards

Portsmouth Harbor Events and Conference Center  
100 Deer Street, Portsmouth, NH 03801

September 28-29, 2017

#RILS2017 @NCIEA1

### Conference Summary

**Scott F. Marion, Executive Director, Center for Assessment**

The Center for Assessment's conference on Assessing Student Learning of the Next Generation Science Standards (NGSS) brought together science and assessment experts and practitioners to wrestle with the immense challenges of creating next generation state science assessments.

The conference agenda essentially traced the journey of the test development cycle. We started with a look at the 3-dimensional nature of the standards and the learning opportunities afforded students by the NGSS led by Bill Penuel of the University of Colorado Boulder and TJ Smolek of the Michigan Department of Education. Next, content and assessment experts (Brian Gong, Center for Assessment. Kathleen Scalise, University of Oregon), a state assessment leader (April McCrae, Delaware Department of Education), and an assessment vendor (Kevin King, WestEd) explored the challenges of articulating the key claims that states want to make about students' science learning. This is particularly tough with the NGSS because the learning targets are so broad, especially given the expansive realm of possible intersections of the multiple dimensions of the standards. The "sparse" nature of the ways in which state assessment designers will have to sample the standards to create an assessment was a theme that ran throughout the conference. Building on this theme, conference participants examined the implications for score reports. As noted measurement expert Ron Hambleton likes to say, "score reports are the only ways that we communicate with the public and the last thing we attend to in the assessment design process." At this year's RILS, Nathan Dadey, Center Associate, along with Jon Cohen (AIR), Kevin King (WestEd), and TJ Smolek (MI DOE), moved us into a discussion of the importance of reporting design early in the design cycle to help clarify what we want assessments to tell us.

One of many highlights of the conference was an “item walk in which 7 different assessment vendors and researchers offered conference participants that chance to do some window shopping to experience a range of test items and tasks designed to assess student learning according to the 3-dimensional standards.

After experiencing the items, we moved to one of the most vexing issues facing the design of next generation science assessment. Alignment is the promissory note that ensures that what is supposed to be taught (the standards) is represented fairly on the assessment. Alignment has been considered a deal-breaking criterion for U.S. Department of Education assessment evaluations, so it is critically important for states to learn how to meet these criteria while designing a practical and useful assessment system. The sparse sampling framework creates challenges for assessments to meet traditional alignment criteria. Our experts—Aneesha Badrinarayan (Achieve), Chris Harris (SRI), and Sara Christopherson (University of WI)—outlined the challenges but also described ways in which alignment processes and criteria might need to change to more fairly represent assessments of 3-dimensional science learning. Our panel of state leaders from Michigan (TJ Smolek), Louisiana (Jan Sibley), Kentucky (Sean Elkins), and Delaware (April McCrae) described how they are facing these challenges and approaching solutions.

The second morning began with Joseph Martineau and Leslie Keng, both Senior Associates at the Center for Assessment, along with Jon Cohen (AIR) and Kathleen Scalise (University of Oregon), getting into the weeds of educational measurement. Most of the attention associated with assessments of 3-dimensional science standards has focused on creating high-quality items and tasks and thinking about alignment hurdles. However, everyone designing a summative assessment cares a great deal (or will soon!) about whether a 2018 score supports the same or at least very similar inferences about what students know and can do who earn that same score2019. This sounds so simple, but it is difficult in the best conditions and the 3-dimensional science standards make this very challenging. A lot of people do not want to think about the intricacies of measurement models and equating designs, but are glad that experts at the Center obsess about such details.

The last session of the conference focused on a subject a lot of people are talking about but have yet to fully articulate. Scott Marion, Center Executive Director, and Bill Penuel, a Professor at the University of Colorado Boulder, were both National Research Council committee members that produced the report, *Developing Assessments for the Next Generation Science Standards*. They described key recommendations in the report that argue for a system of assessments to meet various needs of assessment stakeholders, from teachers and students to state policy leaders. Scott and Bill then laid out some ideas for helping states and districts move toward such a system. They were joined by Sean Elkins (KY DOE) and April McCrae (DE DOE), who described the ways in which their states are implementing what we ended up referring to as “loosely-coupled” systems of assessment.

We finished with a working session where small groups of participants focused on various issues associated with describing what the future of NGSS assessment might look like. Participants emphasized looming questions and how current and future design efforts might address them.

Center Associate Director, Chris Domaleski, brought the conference to a humorous conclusion with a series of internet memes that eloquently summarized what we had learned.

The experts assembled at RILS helped those in attendance share lessons learned and pool thinking to address common concerns. The conference helped us check in on the state of the art of NGSS assessment – providing an overview of development efforts across schools, districts, states and organizations that will serve as a touchstone for future development efforts. Click [here](#) to download the conference presentations, blogs, and papers.