



Research Brief: Designing an Assessment System that Measures Three-Dimensional Science Learning

Summary

Many districts and states need an assessment system that aligns with the [NRC Framework for K-12 Science Education](#) and [NGSS](#) to determine whether students are developing proficiency in science and whether students have adequate opportunities for [“three-dimensional” science learning that blends \(a\) science and engineering practices, \(b\) crosscutting concepts, and \(c\) disciplinary core ideas](#). Assessments serve different purposes, so a system must comprise many different kinds of assessments—but all of them should include tasks that integrate the three dimensions.

BACKGROUND

- “Knowing” science entails more than students demonstrating knowledge of science facts. Students should be able to discuss and make arguments about science, develop and use models, generate and analyze data, and make connections to crosscutting concepts.
- Current assessments use large-scale standardized tests that focus heavily on measuring student knowledge of discipline-specific ideas and concepts. These summative assessments do not consider multiple dimensions of student learning or classroom practices that facilitate student engagement.

Things to Consider

- States face a challenge in determining student proficiency in science. The [NRC Framework](#) and [Next Generation Science Standards \(NGSS\)](#) call for three-dimensional science learning that integrates science and engineering practices, crosscutting concepts, and disciplinary core ideas. States currently rely on one type of large-scale summative assessment to both measure student-level proficiency and school-level performance. Because these monitoring assessments focus on measuring students' knowledge of discipline-specific ideas and concepts, they [fail to attend to multiple dimensions of student learning or classroom practices that facilitate student engagement in science](#).
- Policymakers must ensure that assessment systems are designed to account for the integration of these three dimensions. A broader assessment system of multiple and varied assessments is needed to measure various aspects of learning that collectively, and more accurately, demonstrate science proficiency. A comprehensive assessment system would value internal classroom assessments and student work samples that complement external statewide monitoring of school and district performance.
- [For students to demonstrate proficiency in science, they must be able to meet "performance expectations" at their respective grade levels.](#) These expectations ensure that students are able to engage in science practices that apply crosscutting concepts to disciplinary content knowledge. Standardized assessments should include performance tasks, which students could demonstrate through a collection of work in a portfolio. Thus, designing multiple and varied forms of assessment is key to appropriately determining student proficiency in science and using assessments to improve instruction.
- Traditional statewide assessments do not typically consider classroom practices, school resources, or [teacher professional development opportunities that can impact student proficiency in science](#). An assessment system must audit and attend to equity in opportunities to learn, highlighting systematic differences in resources that inhibit equal access to three-dimensional science learning.



Designing a New Assessment System

- [At all levels, attends to equity and fairness by making the most of students' diverse linguistic and cultural resources and ensuring adequate opportunities for all students to learn science.](#)
- [Starts with developing internal classroom assessments, so that formative assessment is used to inform instructional practices.](#)
- [Includes various types of formative and summative assessments embedded into instructional practices.](#)
- [Monitors effectiveness of science curricula, classroom practices, and opportunities to learn through student and teacher surveys, collected work of students, and observations.](#)
- [Considers the critical role of technology in both student learning and assessment.](#)
- Refrains from attempts to assess every standard in every grade, for every child.
- [Requires a gradual, phased-in implementation that evolves alongside changes to curriculum, instruction, and professional development.](#)
- Implements state monitoring assessments at least once, but no more than twice per grade span (K-5, 6-8, 9-12).
- [Includes performance tasks within external monitoring tasks.](#)
- [Analyzes costs, benefits, and tradeoffs of designing and implementing a new system.](#)

Implications for Practice

To measure the three dimensions of science learning:

- [Assessments must go beyond the traditional types of high-stakes summative assessments that often define standardized tests.](#)
- [Student proficiency should be measured more accurately through multiple, varied forms of assessment that occur at the classroom and state levels.](#)
- [States must adopt a set of standards that align with science and engineering practices, crosscutting concepts, and disciplinary core ideas.](#)
- New assessments require co-design among stakeholders with varied forms of expertise.
- [Co-design teams must consider how to measure both student learning in multiple dimensions and the effectiveness of curricula.](#)
- [Policymakers need a way to address the incorporation of classroom-embedded assessments into systems of monitoring and accountability.](#)

To promote equity and fairness:

- Value the diverse experiences, viewpoints, and cultural and linguistic resources of all students.
- [Connect students' cultural and linguistic resources to disciplinary knowledge and practices.](#)
- [Provide adequate school resources to support student learning.](#)
- [At the state level, monitor individual schools to ensure that all students have access to adequate resources and opportunities for science learning.](#)
- [Use formative assessment to elicit students' diverse interests and experiences in relation to their science learning \(see "Micros and Me" example\).](#)

SAMPLE ASSESSMENTS

- Classroom Formative Assessment of Three-Dimensional Learning During An Investigation: ["What Is Going on Inside Me?," Middle School](#)
- Evidence-Centered Design of a Task Assessment Using Design Patterns: ["Pinball Car Task," Middle School](#)
- Learning Progression on Data Modeling with a Classroom Embedded Task: ["Measuring Silkworms," 3rd grade](#)
- Formative Assessment Analysis of Classroom Discourse: ["Behavior of Air," Middle School](#)
- Formative Assessment of Student Understanding Using Clicker Technology: ["Movement of Water," Middle School](#)
- Using Multiple Interrelated Tasks to Assess a Core Idea Across Multiple Practices: ["Biodiversity in the Schoolyard," Middle School](#)
- Online Task Used for Formative and Summative Assessment: ["Climate Change," High School](#)
- Using Embedded Computer Simulation-based Modules in Curriculum for Formative and Summative Assessment: ["Ecosystems," Middle School](#)
- Mixed Item Format Items in Monitoring Test: ["Photosynthesis and Plant Evolution," High School, AP Biology Exam Example](#)
- Swiss Example of Three-Dimensional Learning Performance Assessment Item Used for Monitoring: ["Floating and Sinking," 2nd grade](#)
- Performance Task Used in the Classroom for Curricular Evaluation: ["Plate Tectonics," Middle School](#)

REFLECTION QUESTIONS

- What are science assessments currently like in your state? What is missing from your state's science assessment system?
- How does your state currently determine student proficiency in science? What are the science "performance expectations" within your state and how do those expectations align with NGSS?
- Who should be involved in designing and implementing a three-dimensional assessment system for science learning?
- How would you design an assessment system that accounts for science and engineering practices, crosscutting concepts, and disciplinary core ideas? In this system, how would multiple and varied forms of assessment be designed both to evaluate student proficiency in science and to use assessments to improve instruction? How would this assessment system link internal classroom assessments and external monitoring systems? See [Chapter 4 and 5 in *Developing Assessments for the Next Generation Science Standards*](#).
- How can a new assessment system attend to equity in science education? See [NGSS, Appendix D, "All Standards, All Students"](#) and [NRC Framework, Chapter 11](#).
- What teacher professional development models would you propose to build and refine formative classroom assessments?
- What are the challenges and opportunities your state faces in adopting an assessment system for three-dimensional science learning? See [Pellegrino, 2013](#).



References

- National Research Council. (2014). [*Developing Assessments for the Next Generation Science Standards*](#). Washington, DC: National Academies Press.
- National Research Council. (2012). [*A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*](#). Washington, DC: National Research Council.
- National Research Council. (2007). [*Taking science to school: Learning and teaching science in grades K-8*](#). Washington, DC: National Academies Press.
- NGSS Lead States. (2013). [*Next generation science standards: For states, by states*](#). Washington, DC: National Academies Press.
- Pellegrino, J. W. (2013). [*Proficiency in science: Assessment challenges and opportunities*](#). *Science*, 340, 320-323.

ALSO SEE STEM TEACHING TOOLS:

- #4 [Multiple Instructional Models](#)
- #14 [Why NGSS?](#)
- #15 [Equity in Science Education](#)



School of Education
UNIVERSITY OF COLORADO BOULDER