The Next Generation Science Standards (NGSS) identify what students should know and be able to do across three dimensions: scientific and engineering practices, disciplinary core ideas, and cross-cutting concepts. No single assessment or current statewide assessment system is adequate to the task of measuring whether students have met the NGSS. As has become apparent to educators and policymakers, a different approach is needed.

Researchers and practitioners suggest a comprehensive approach that both monitors what students have learned—the assessment type with which most people are familiar—and serves as a diagnostic tool to improve instruction. This second practice is new for states but common in many other countries. Such assessment becomes a professional development tool for teachers, informing them of the effectiveness of their instructional approaches and helping them identify ways to modify these approaches to get better results.

A comprehensive assessment system meets multiple goals and addresses different audiences’ needs. Administrators and policymakers need to monitor student performance. Teachers need information that enables quick changes in instruction and curriculum.

Current assessments are not comprehensive, nor do they measure the learning that the NGSS describes. New assessments, consistent with the NGSS, will necessitate planning and investment. While this may seem challenging, it is possible. Such a system should include “multiple components that reflect the connected use of different scientific practices in the context of interconnected disciplinary ideas and cross-cutting concepts, address the progressive nature of learning by providing information about where students fall on that continuum, and include an interpretive system for evaluating a range of student products that are specific enough to be useful for helping teachers understand the range of student responses and tools for deciding the next steps.”

The good news is that there are examples for designing comprehensive assessments (e.g., box 1). For example, many teachers conduct assessments that include classroom observations, student projects, quizzes, and exams. If coordinated and normed, these assessments can help teachers gauge what students know for the purpose of changing instruction, materials, or strategies, as well as providing data for use throughout the school. Multiple districts have developed assessments through research projects that show it is possible to use classroom assessment to measure the three dimensions of the NGSS. A number of states had statewide performance assessments in the late 1980s and early 1990s that were designed exactly for these purposes—monitoring and classroom information—but they were discontinued because of some technical problems, time, and the No Child Left Behind (NCLB) requirements. Nonetheless, educators learned much from that past work.

Today there are good examples of high-quality assessment for monitoring student performance: the National Assessment of Educational Progress (NAEP), the New England Common Assessment Program (NECAP), and the statewide assessment in New York for grades 4 and 8. These are mixed-item formats with summative and formative tasks. The two current multistate consortium assessments are to some degree examples for mathematics and English Language arts, but they were designed to meet the technical requirements of NCLB rather than the newer vision of multiple dimensions and real-time classroom information.

Policy Considerations

State boards of education have many things to consider when putting an assessment system in place. Key considerations are who will use the data, what should be measured, and the timeline for implementing new assessments. Other considerations for states, not dealt with here, include the extent of student opportunities for science learning, professional capacity to administer the assessments, cost, alignment with other policies, reliability of the assessment tools, and equity issues such as culture, language, ethnicity, gender, and disabilities.

Who will use the assessment data? Whatever the assessment system, robust reporting is key and requires a deep sense of the purposes of the assessment and what various audiences will do with the collected data. The information that emerges from an assessment must be in a form...
that can support performance monitoring, students’ application of knowledge, instructional improvement, and teachers’ professional growth.

**What do you want to measure?** Traditional multiple-choice tests and computer-generated assessments have limits in measuring what students know. A multileveled assessment is needed to capture helpful, useable information and is essential for measuring NGSS performance.

**What is the timeline for developing and implementing assessments?** Go slow with the rollout. A planned, deliberate process will allow for building a robust collection and reporting system and still be able to address shorter term requirements for monitoring and classroom assessment. Time is also needed to train teachers in use of the assessments, in analysis of the data, and to consider how data are to be shared with the public.

**Conclusion**

Much is known about science learning, the materials to support that learning, and how to assess that learning. Most of what is recommended here is already being done in some schools. State boards and state departments of education can develop, review, and revise their assessment systems as they implement these new science standards. Their task is to make the appropriate adjustments that enable schools to meet the expectations for determining what our students know and can do in science. Development of the assessments, professional development, and the implementation of new technologies will take time but are worth the effort. The new model for assessment represents the next step in sophistication for better learning and better measurement.

_Francis Eberle is NASBE’s deputy executive director, francise@nasbe.org. This publication was made possible by the Carnegie Corporation of New York._

---

**Box 1. Biodiversity in the School Yard: Sample Assessment Tasks for Grade 5**

| Task 1: Collect data on the number of animals and species in school yard zones. | —Provides the teacher information about students’ ability to collect and record data for other tasks. |
| Task 2: Create bar graphs that illustrate abundance and richness patterns in the data from each school yard zone. | —Provides information about students’ ability to display data and whether students understand the core concepts of species abundance and species richness. |
| Task 3: Which zone has the greatest biodiversity? Construct an explanation to support your answer. | —Hints and supports are provided to help students make claims and support them. |
| Task 4: Which zone has the greatest biodiversity? Construct an explanation to support your answer. | —Teacher provides class data but this time with no hints or supports. Provides teacher information about whether students understand the core concepts and what constitutes a full explanation. |


---

**Resources**


---

**Notes**


5. National Research Council, _Developing Assessment_.

6. Ibid., _Conley and Darling-Hammond, Creating Systems_.

7. National Research Council, _Developing Assessment_.

8. One example is the Maryland School Performance Assessment system.