Since the release of the NGSS, science classrooms across the United States have shifted science learning away from a focus on decontextualized facts and concepts and toward engaging students in making sense of the world around them. Many educators have looked to assessments to provide feedback about students’ progress in meeting this new vision of science teaching and learning, but few assessments sufficiently reflect NGSS goals (Wertheim et al. 2016). The National Research Council (NRC) identified performance assessments as an essential tool for guiding shifts toward three-dimensional (3D) learning (NRC 2014), though many questions remain about how to design assessments that provide the information that teachers need to guide their work.

The Stanford NGSS Assessment Project (SNAP) was established in 2015 to address these questions and has developed a particular focus on how to design and use performance assessments as part of an equity-seeking system of supports for the NGSS. SNAP responded to a groundswell of demand for supports for 3D assessment by creating exemplar assessments, step-by-step design tools, and professional learning modules for using 3D performance assessments to center student learning. Our assessments are designed to be a continuation of teaching and learning;...
students draw on the competencies they are developing during instruction in new ways as they make sense of real, meaningful questions and problems (Wertheim 2021). Each SNAP assessment presents a complex task but provides scaffolds that create entry points for all students to share their thinking. By focusing on the strength of students’ scientific reasoning instead of right or wrong answers, the tasks are designed to invite students working at any level of proficiency to respond and to value the diverse ideas, interests, and perspectives that students bring to real phenomena. These features both drive students’ motivation to engage and persist through complex, multipart assessments and create opportunities for them to use their experiences as resources that can assist their sensemaking.

Simply improving access to high-quality 3D performance assessments, however, is not transformative. If assessments are to guide shifts in teaching and learning, they need to be embedded in growth-oriented instructional practice. This means using students’ responses to tailor instruction around their strengths and needs and inviting students into this process as agents of their own learning (Black and Wiliam 1998). When teachers use products from performance assessments to construct frequent, descriptive feedback, engage students in reflecting on their own progress, and build from students’ ideas and experiences, they shift from simply evaluating students toward driving deeper learning (Hattie and Timperley 2007). Making this transition requires a change in how assessments are used. Instead of functioning primarily as instruments for grading, assessments become part of a dynamic and responsive instructional practice oriented toward ensuring all students have the supports they need to continuously develop their 3D learning.

3D performance assessment in a classroom

This vision of assessment often raises questions from teachers about what it can look like in practice. This article describes a sample SNAP assessment, paired with a classroom vignette, to illustrate how a performance assessment can be used to inform equitable, efficient, and effective instructional practices. The SNAP short performance assessment (SPA) Will Communities Run Out of Meltwater? (referred to as the “Meltwater SPA”) is a 40-minute task that is intended to be completed individually after instruction on the focal performance expectation (HS-ESS2-2). This vignette describes one way a teacher uses these assessments to drive student-centered and equity-seeking instruction. This teacher is fictional, but we use her to illustrate a compilation of exemplar assessment practices we have observed in real science classrooms.

Ms. Wong teaches a ninth-grade physics class. To address performance expectation HS-ESS2-2, she has students analyze data and use computational models to explore Earth system dynamics. Early in this unit, she introduces a student-facing rubric (Figure 1) for the Meltwater SPA, which is her end-of-unit assessment. The rubric describes the range of ways, from emerging to excelling, students use the three dimensions during the culminating question for the SPA. Students discuss in pairs and restate each level in their own words to ensure that they understand the criteria. At several points during the unit, students use this rubric to evaluate their own progress, look to the next levels, and identify ways to improve their work.

Introducing this rubric early in the unit and revisiting it to reflect on their progress makes the connection between instruc-

FIGURE 1

Student-facing rubric aligned to the culminating question for the assessment for HS-ESS2-2.
tion and growth-oriented assessment explicit for students. The rubric also communicates priorities for learning to students: They are being evaluated on the strength of their reasoning with evidence, not their knowledge of specific facts or terms.

By the end of the unit, Ms. Wong’s students have practiced using the dimensions in a variety of ways to make sense of their unit’s anchoring phenomenon. The Meltwater SPA provides an opportunity for her students to transfer the competencies they have been developing to a new phenomenon (Figure 2).

Efficiently assessing student work
Performance assessments are often characterized as summative tests that are time-intensive to grade. But by employing strategies for using the assessment for learning, they can inform ongoing growth effectively and efficiently, even when they are used at the end of a unit. Performance assessments often pose a driving question or problem about a phenomenon that students resolve at the end of the task, and students make visible their progress with the three NGSS dimensions in various combinations and modes as they work through the task. This format builds toward the final, culminating prompts such that they often provide the richest information about students’ progress with using the three dimensions. For efficiency, teachers frequently begin their analysis by examining a small sample of student work for only the final prompts to identify trends that roughly represent the variation across their class. If the trends in students’ use of the three dimensions are difficult to discern, the earlier prompts can be used as sources of additional evidence.

Ms. Wong first focuses her attention on the culminating question of the assessment, Question 5 (Figure 3), because it demonstrates how students use the grade-level expectations for all three dimensions to resolve the assessment’s driving question. She analyzes a small sample of student work using her rubric to look for two types of trends: (1) common strengths and needs with the three dimensions across her entire class; and (2) common trends among students who are working at the earliest stages of proficiency with this PE.

She sees that most of her students tend to make a relevant claim about the phenomenon and that they support their claim with interpretations of the data (SEP) and some reasoning about the interactions between Earth systems (DCI) (see example in Figure 4). This shows her where their strengths are and also where there is room to grow. Very few of her students are using sufficient reasoning about climate feedback effects, particularly how they affect the stability of the climate system (CCC). Their thinking about feedback effects remains implicit and so they are not yet providing a complete explanation for how the data support their claim.

FIGURE 2
Introduction to the phenomenon of the “Will Communities Run Out of Meltwater?” assessment.

Introduction
The river basins around the Himalayan Mountain Range are some of the world’s most densely populated areas (see map). There are over 900 million people in the region, and they rely heavily on water from glaciers and snow from the mountains (meltwater) for drinking and to grow crops.

In recent years, there has been a huge increase in meltwater. Local communities are trying to figure out: Is this just part of a natural cycle and the meltwater will return to normal levels in future years? Or should they be worried that they will run out of meltwater in the near future?

In this task, you will analyze geoscience data to help local communities investigate the problem so they can better prepare for the future.

FIGURE 3
Culminating question of the “Will Communities Run Out of Meltwater?” assessment, which assesses all three dimensions of the PE.

5. Based on your flowchart and your data analysis, should communities near the Himalayan mountains be worried that they will run out of meltwater in the future? Why or why not? In your explanation include the following:

• Explain how increases in meltwater cause feedback effects.
• Discuss whether feedback effects stabilize or destabilize the climate system.
• Cite evidence from multiple figures, the computational model, and your flowchart model to support your explanation.
Sample student response to Question 5, color-coded for evidence of where they are using three dimensions.

Based on my flowchart and data analysis, communities near the Himalayan mountains should be worried that they will run out of meltwater in the future because an increase in meltwater destabilizes the climate system. We can see in Figure 1 that because of the increase in meltwater, the albedo on Earth has decreased. Albedo plays a significant factor in the global air temperature, and so if the albedo decreases, the Earth becomes warmer. This is shown in Figure 3 where we see the increase of global mean surface temperature. It is also shown in Figure 4 where we see the Snow Cover Extent and Sea Ice Area decreasing because of lower levels of albedo from global warming.

Their use of the SEP (analyzing and interpreting data) is indicated in blue, the DCIs (Earth materials and systems and weather and climate) in orange, and the CCC (stability and change) in green.

In addition to considering trends across the class, analysis of student work should also bring attention to the specific needs of students working at the emerging level.

The first two questions of the Meltwater task (Figure 5) are scaffolds for students; they are intended only to deepen students’ understanding of the significance of the phenomenon and focus their attention on analyzing each graph before they need to use the data to reason about the phenomenon. These scaffolds may offer the clearest evidence of some students’ proficiency with the PE. Students who struggle with the demands of the more complex prompts may provide responses to Question 5 that are difficult to interpret or may leave it blank. Questions 1 and 2 may offer crucial insights into these students’ progress toward the PE.

Ms. Wong selects a few samples of student work at the emerging level and reviews their responses to Questions 1 and 2. These prompts elicit evidence of just two dimensions together and are less complex than the later questions. In her analysis, it becomes clear that most students understood the problem they were investigating and were successfully extracting relevant information from some of the graphs (SEP). However, they were struggling to combine and incorporate data meaningfully into their reasoning about the phenomenon in later questions (SEP). This observation helps her pinpoint precisely what the students are making progress on and where these students are ready to focus on growing.

Using assessment data to guide effective and equitable instruction

Performance assessments provide crucial insight into students’ progress with 3D learning, but assessments, even end-of-unit assessments, have the potential to do more than simply evaluate students’ learning. Efficient and effective instruction uses evidence from all assessments to craft moves that address the specific strengths and needs of the students (Heritage and Wylie 2018). For this assessment practice to be equitable, teachers must ensure that small groups of students who are at earlier levels of proficiency have the supports they need to benefit from the moves targeted at the majority of the class. As an efficient next move, a teacher might ask students to self-assess their work relative to one of the whole-class trends or the student-facing rubric and to explain how they improved their responses. A teacher might also look to the next units where the same SEPs or CCCs are used and provide supports tailored to students’ needs. Alternatively, a teacher might insert brief instructional activities immediately following the assessment that attend to these trends and the ideas, interests, and experiences that students reveal. Which move, or combination of moves, a teacher makes depends on the nature of the trends, their prevalence, and instructional time.

Before Ms. Wong evaluates each of her students’ assessments in depth, she has them work in pairs to practice strengthening the reasoning about climate feedbacks in a sample response. Then she has them turn to their own work to do the same individually. Students use the student-facing rubric (Figure 1) to self-assess, write reflections about how they improved their response to meet a higher level of the rubric, and describe what they plan to do in the future to continue this growth. She reads the students’ revised responses and self-evaluation and if she agrees with it, she provides a grade. If she does not agree, she adds feedback and students revise again before she adds a grade.

For students working at the emerging and developing levels, she identifies opportunities in the next unit to keep practicing incorporating data into their explanations in a new content area. She plans an additional set of scaffolds for an activity early in the next unit to provide targeted support for these students, and she develops exit tickets focused on this practice to monitor their progress.

Ms. Wong’s story exemplifies many of the transformations in how teachers are using assessments that we see across the country. One teacher who worked with SNAP to incorporate these shifts into her assessment practice reflected on the impacts: “The work that we’ve been doing in performance assessment... is firmly seated in supporting teachers’ knowledge of students, but it’s really also effective at engaging students in their own learning. It ensures that we as their teachers are being really thoughtful about who our students are and what we want them to do and ensuring our instruction gets them there.”
Excerpts from scaffolding questions (1 & 2) from the “Will Communities Run Out of Meltwater?” assessment that are not intended to be formally assessed.

1. First, investigate what is happening to the sources of meltwater in the Himalayas. Use the data and the questions below to guide your analysis:

   a) What has happened to the glaciers in the Himalayas in recent years? Use evidence from Figure 1 to support your response.

   b) What has happened to snow cover in the Himalayas in recent years? Use evidence from Figure 2 to support your response.

2. Now that you have studied data showing changes to the glaciers and snow cover that are the sources of meltwater in the Himalayas, examine data that will help you learn about the recent increase in meltwater. Analyze each of the graphs below and record your analysis in the right-hand column.

<table>
<thead>
<tr>
<th>DATA SOURCE</th>
<th>YOUR ANALYSIS: WHAT DOES THIS GRAPH TELL YOU?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snow Cover Extent and Sea Ice Area, 1980 - 2020</td>
<td></td>
</tr>
</tbody>
</table>

![Graph showing the amount of annual Arctic sea ice and snow cover over time.](image1)

*Figure 3. Graph showing the amount of annual Arctic sea ice and snow cover over time.*

<table>
<thead>
<tr>
<th>Global Mean Surface Temperature, 1880 - 2000</th>
</tr>
</thead>
</table>

![Graph showing changes in the annual mean temperature of the surface of the Earth over time.](image2)

*Figure 4. This graph shows changes in the annual mean temperature of the surface of the Earth over time.*
Supports created: SNAP tools

Performance assessments offer the most power for transforming science teaching and learning when they are part of a system in which teachers collaborate around the intersection of instruction and assessment. Just as sensemaking in science thrives within a culture of discourse and debate among students, use of complex assessments to drive instructional decisions flourishes in a school culture where colleagues have the opportunity to discuss, challenge one another’s thinking, and plan strategic next steps together. To build this culture, however, assessments must be viewed not merely as a vehicle for grading but instead as a key to centering students in instruction. However, few schools or districts dedicate the time and resources to the professional learning needed to build capacity to make these shifts.

The combination of the challenge and promise that transitioning to the NGSS presents has led teachers to seek resources that will help them gain actionable insights into their students’ learning. SNAP’s free tools and hybrid courses meet this demand by offering teams step-by-step guidance for the development and/or use of 3D performance assessments to center student learning. The courses attend to both the limited time and resources dedicated to science assessment and the value of approaching classroom assessment as part of a system. They are self-paced and are completed in teams that may be composed of teachers at any grade level, school and district administrators, instructional coaches, and others. The teams cultivate a community of practice around the use of assessment for learning within a school or district.

Looking forward

The answer to the question of what it means to assess the NGSS efficiently, effectively, and equitably only begins with the assessment itself. The next step is a shift in how these assessments are used, such that students see the assessment as an opportunity to track their progress and chart a course forward, teachers see it as a way to illuminate students’ assets as well as their needs, and administrators see it as a tool for facilitating deeper learning and building coherence across classes, grades, and subjects. As one teacher described the transformation in their assessment practice: “I’m excited for these tasks. I don’t hate assessment anymore! ... And the kids don’t hate it anymore, too!”

SNAP was originally a project of the Stanford Center for Assessment, Learning, and Equity (SCALE). SCALE sunsetted in December, 2021 and the science group moved to WestEd. All of the SNAP resources are available from SCALE Science at WestEd.

ONLINE CONNECTIONS

To see the full Melwater Task visit: https://scalescience.wested.org/snap/snap-assessments/short-performance-assessments/

To learn more about SNAP performance assessments, toolkits, and free courses, visit

- Performance assessments: https://scalescience.wested.org/snap/snap-assessments
- Design tools: https://scalescience.wested.org/snap/ngss-assessment-design-and-analysis-resources
- Courses: https://scalescience.wested.org/snap/resources/snap-courses

REFERENCES


Lauren Stoll (laurenkstoll@gmail.com) is the Director of Professional Learning for SCALE Science and Jill Wertheim (jwerthe@wested.org) is the Director of SCALE Science at WestEd, Alameda, CA.