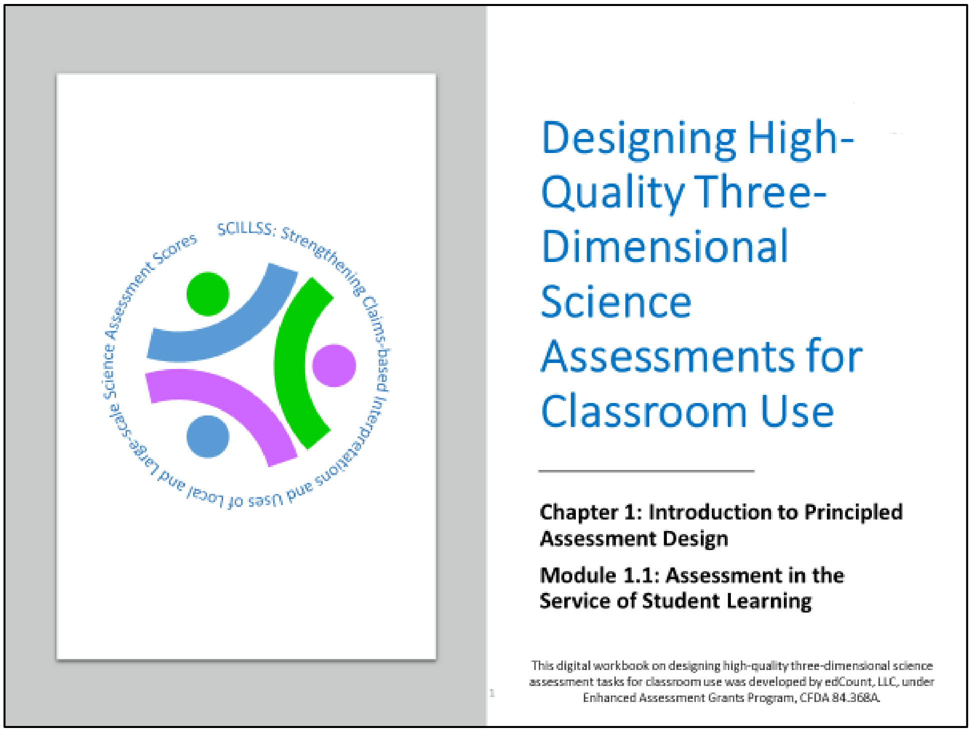
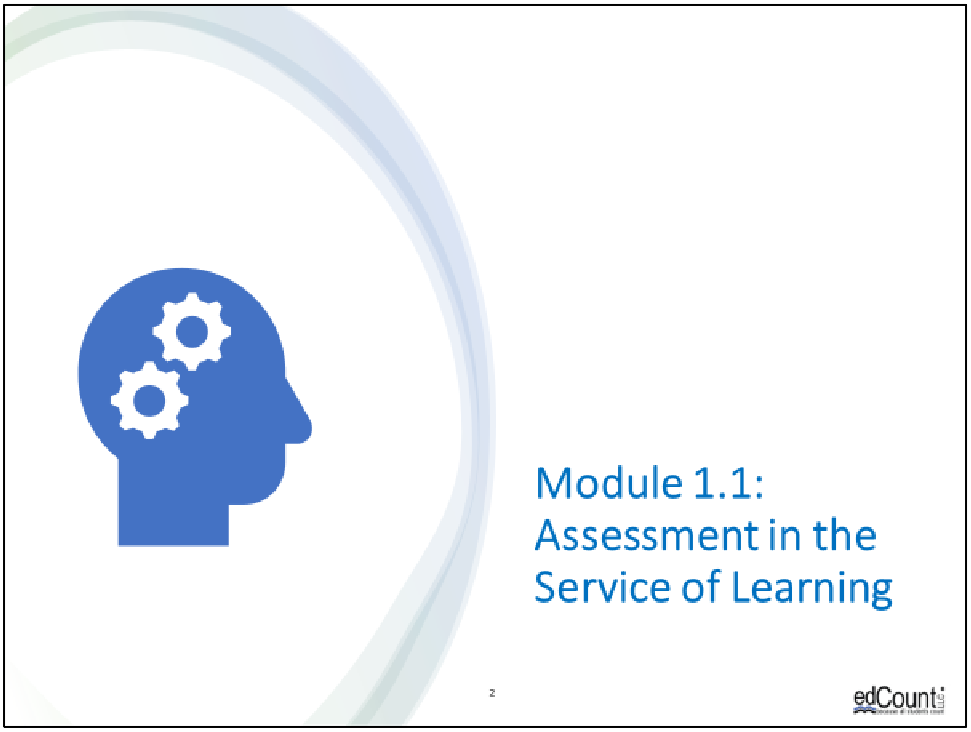


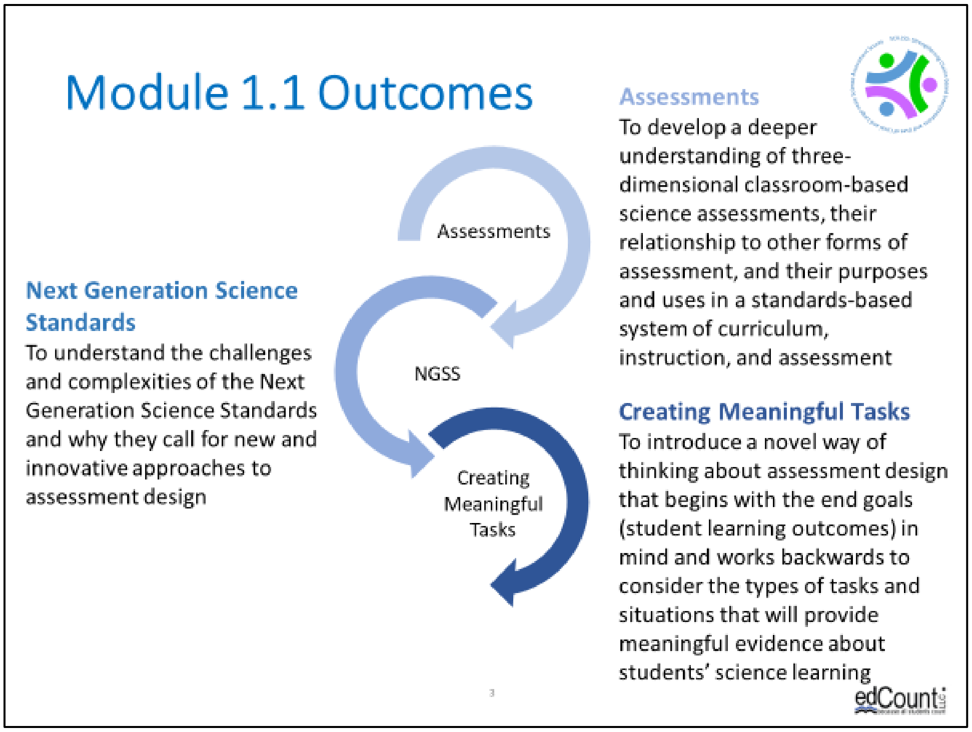
Welcome to the first of four chapters in a digital workbook on designing high-quality three-dimensional science assessment tasks for classroom use. This workbook is intended to help educators design and evaluate tasks that provide meaningful information about what students know and can do in science.

This digital workbook was developed by edCount, LLC, under the US Department of Education's Enhanced Assessment Grants Program, CFDA 84.368A.



Chapter 1 of this workbook includes a series of three modules. Module 1.1 focuses on classroom-based assessments, their relationship to other forms of assessment, and their purposes and uses in a standards-based system of curriculum, instruction, and assessment. Modules 1.2 and 1.3 introduce principled-assessment design, a backward design approach for developing three-dimensional tasks aligned to the Next Generation Science Standards (NGSS) for use within classrooms.

In these Chapter 1 modules, we lay the groundwork for educators to engage in intentional assessment design by considering the purposes and uses of classroom assessments and the types of tasks and situations that elicit evidence that supports meaningful interpretations of students' science learning.

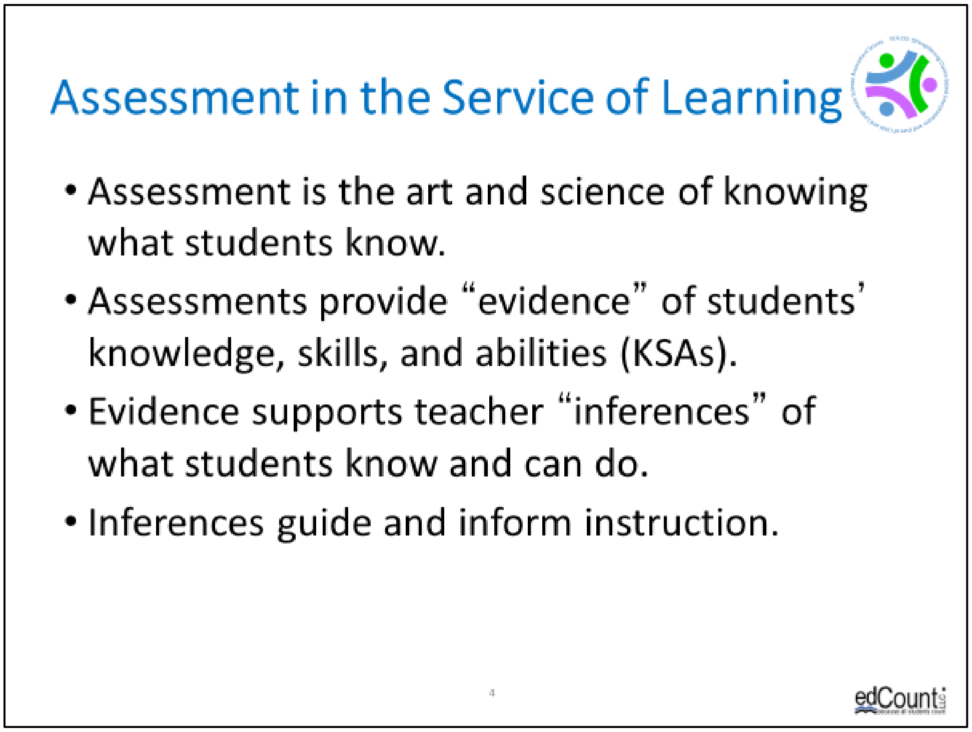


Let's take a closer look at the outcomes associated with Module 1.1. Assessment in the Service of Learning.

Teachers' assessment of student learning—what they know and can do—is integral to teaching every day. This module begins with an introduction to the purposes and uses of three-dimensional science classroom-based assessments, their relationship to other forms of assessment, and their role in a standards-based system of curriculum, instruction, and assessment.

In this module, we also explore the challenges and complexities of the Next Generation Science Standards, or NGSS, from which many state standards are based, and why they call for new and innovative approaches to assessment design.

Finally, we introduce a novel way of thinking about assessment design that begins with the end goals, or student learning outcomes, in mind, and works backward to consider the types of tasks and situations that will provide meaningful evidence about students' learning of three-dimensional science standards.

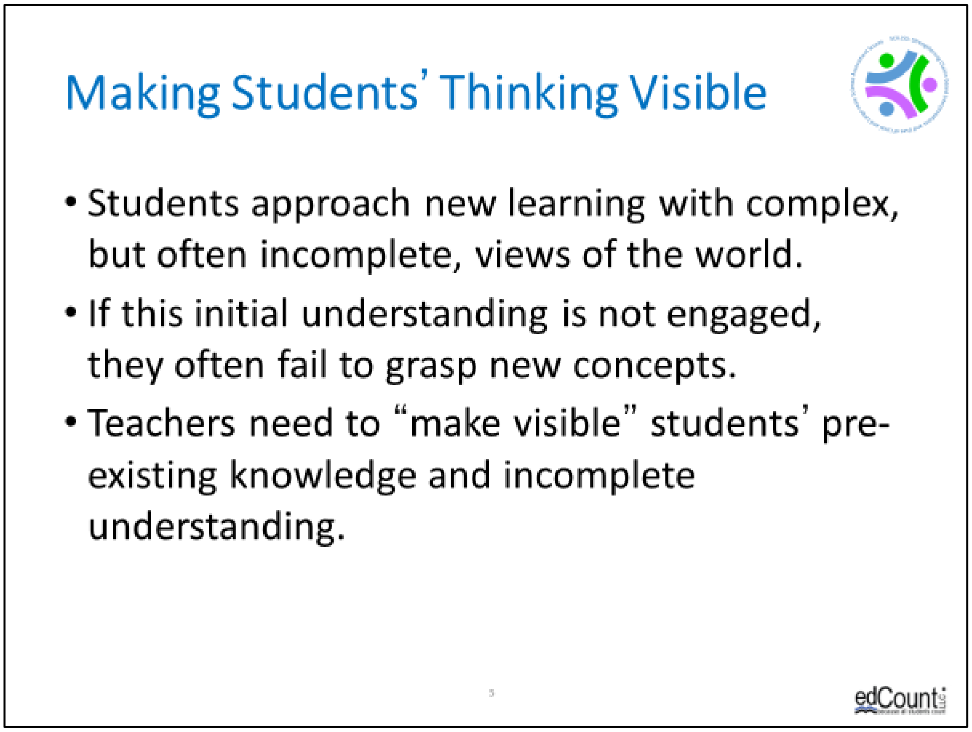


Assessment as the art and science of knowing what students know is a radically new idea. We must begin thinking about assessment design and uses in the service of learning. To do this well, we need to understand the role of assessments in the curriculum and instruction system.

As educators, we assess to guide and inform instruction; if we find out through the process of instruction that students aren't learning key concepts or ideas, then we need to go back and help them develop that understanding.

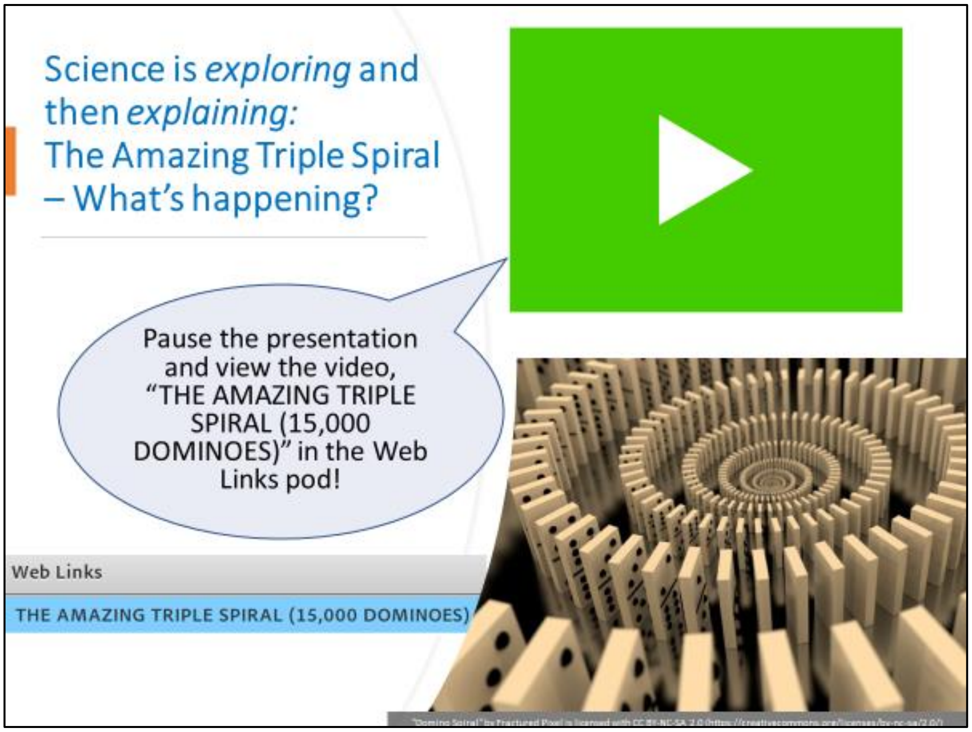
Assessments must also help us recognize how students develop their understanding. We cannot see what our students are thinking. Therefore, we depend upon behavioral evidence—what they say, do, and produce—to give us information about internal processes. In this way, we can make students' thinking visible in order to identify misconceptions and to address them through instruction.

When we give an assessment, what we are looking for as a result of the activity is evidence of students' knowledge, skills, and abilities. By eliciting and interpreting this evidence, we can make inferences about the depth of students' understanding of the assessed topic and make judgments or inferences that guide and inform our instruction.



Students approach new learning with complex, but often incomplete, views of the world. To help students grasp new concepts and build more sophisticated understandings over time, we must acknowledge that their current views of the world may be incomplete and provide opportunities for students to engage in new learning of complex science ideas.

When students' thinking is made visible through teaching, learning, and assessing, teachers can begin to better understand students' previous knowledge, depth of understanding, areas of incomplete understanding, and potential misconceptions.



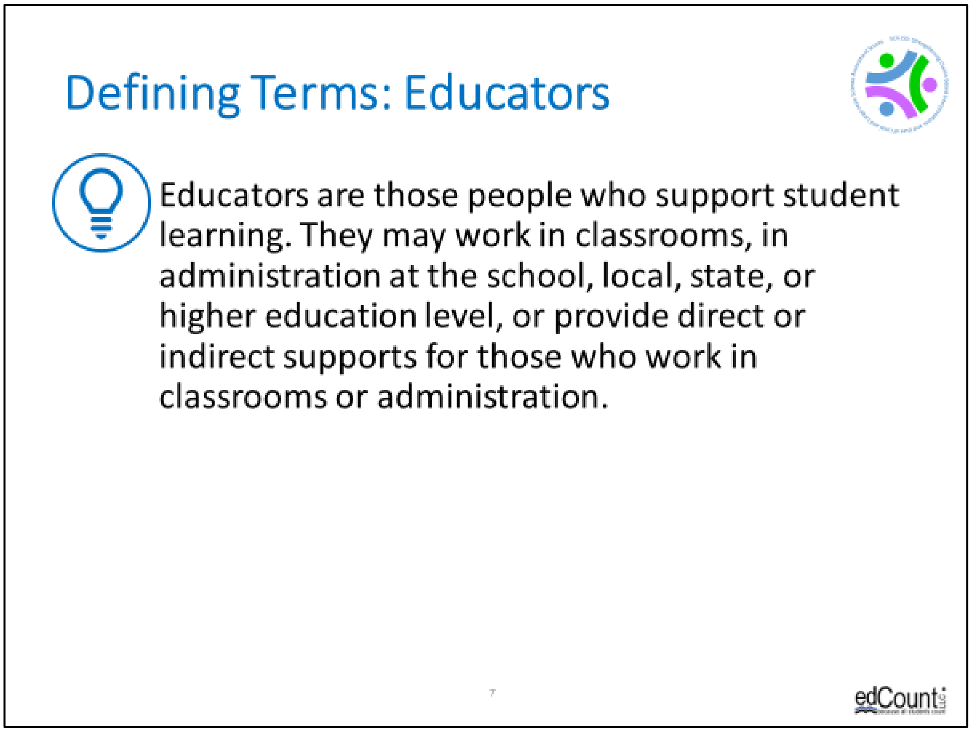
The learning of science is similar for all learners, whether students or scientists or engineers; in order to learn science, you need to do the work of science. Science is both a body of knowledge and an engagement in scientific practice to develop and gradually increase understanding of science and the world we live in.

Our students are naturally curious—they love to explore and ask why. Through exploration and problem solving, students start developing ideas about how the world works. In the science classroom, a carefully chosen phenomenon or engineering design problem can drive student inquiry and exploration. Each can serve as an anchor for exploring science.

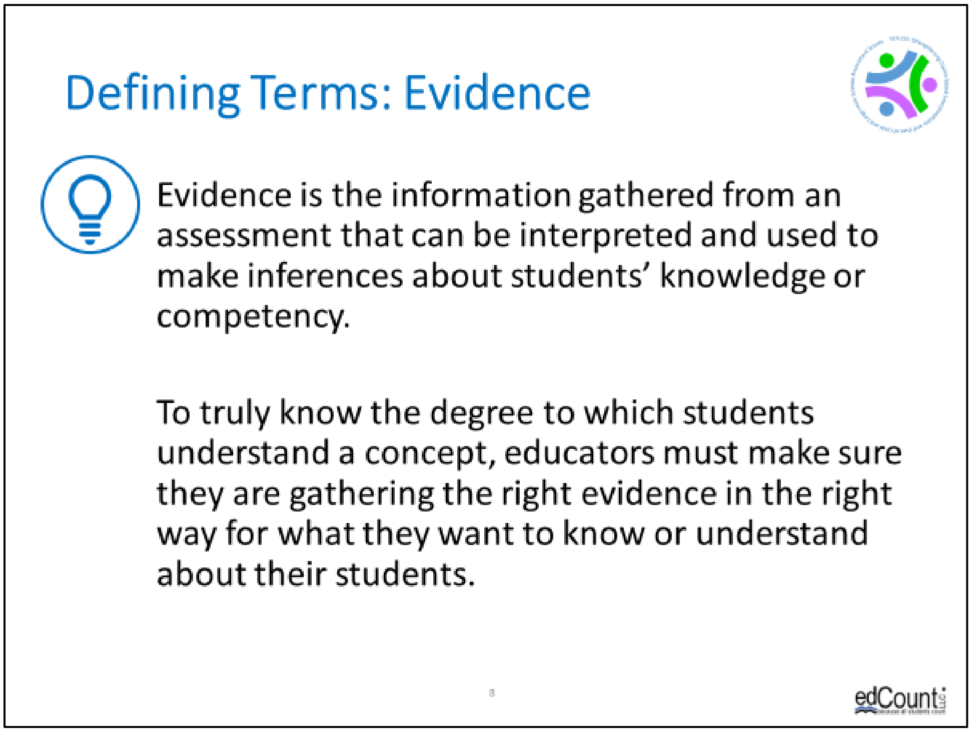
Simply defined, natural phenomena are observable events that occur in the universe. Engineering design problems are born out of meeting human needs and wants and require refining and designing solutions, such as creating objects, processes, and systems, to address those needs and wants. Students can use their science knowledge to explain or predict phenomena or provide possible solutions to engineering design problems—just the way scientists and engineers would!

The phenomenon or design problem is selected based on the topic of the standard and intended outcomes of instruction, and should be observable, interesting, complex, and relevant. This careful selection helps students make sense of the science and their ability to explain it in the context of the phenomenon or design problem.

This short video, "The Amazing Triple Spiral," is an example of a phenomenon that could invoke the question, "What's happening and why?" Good questioning leads to purposeful exploration—the doing of science—which leads to our common goal of increasing students' understanding of science. Please pause the presentation to view the short video.

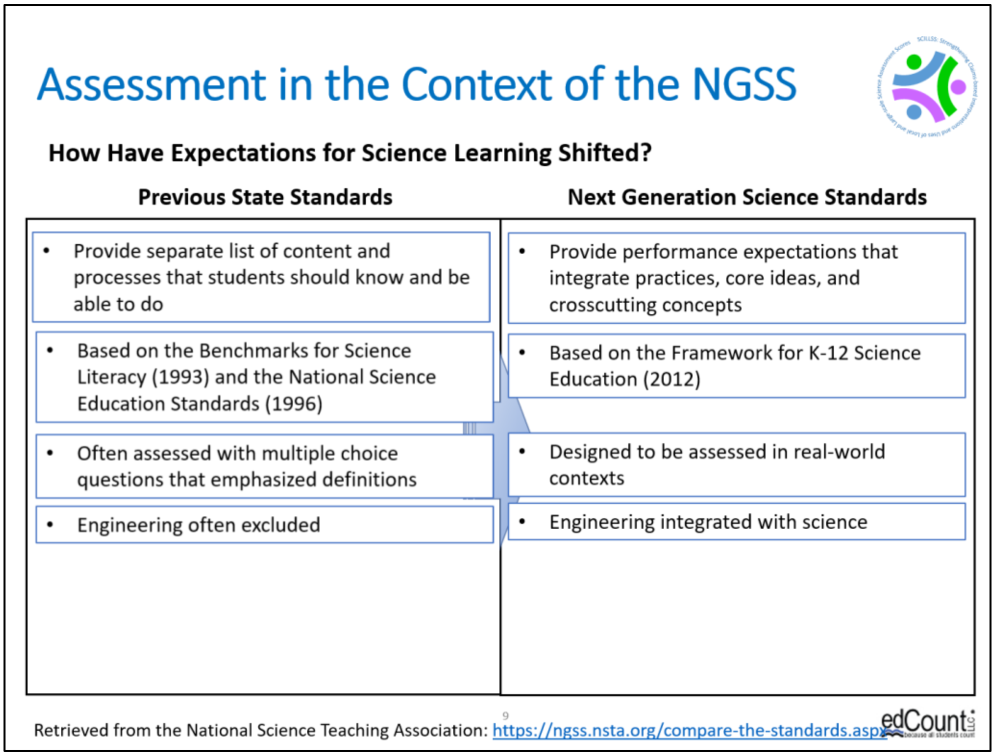


Throughout the chapters, we will clarify the terms we'll be using. First, when we refer to "educators," we mean people who support student learning. They may work in classrooms, in administration at the school, local, state, or higher education level, or provide direct or indirect supports for those who work in classrooms or administration. Note that the "lightbulb" icon indicates a term that is found in the digital glossary to promote a fuller understanding of the presented content.



We will use the term "evidence" to describe the information gathered from an assessment that can be interpreted and used to make judgments about students' knowledge or competency.

To truly know the degree to which students understand a concept, educators must make sure they are gathering the right evidence in the right way for what they want to know or understand about their students.



As educators, what is being taught and our expectations for science learning have shifted. Take a moment to reflect on your previous state science standards compared to the Next Generation Science Standards or your current science standards or indicators. Where have the shifts in teaching and learning occurred?

In this activity, we'll sort various statements to "fill in the chart" to clarify these shifts and point to the call for science educators to have a deep understanding of these three-dimensional standards and implications for teaching and learning.

Let's consider the first two statements:

- Provide performance expectations that integrate practices, core ideas, and crosscutting concepts. Previous state standards or NGSS?
- Provide separate list of content and processes that students should know and be able to do. Previous state standards or NGSS?

Separate will no longer work. In the service of deeper understanding and more sophisticated science learning, instruction must be based on standards that promote the integration and application of knowledge of disciplinary core ideas and crosscutting concepts while engaging in a science or engineering practice.

Here are the next two statements:

- Based on the Benchmarks for Science Literacy (1993) and the National Science Education Standards (1996). Previous state standards or NGSS?

- Based on the Framework for K-12 Science Education. Previous science standards or NGSS?

Clearly, the dates are telling. But the vision of science learning in the National Research Council (NRC) foundational report, often referred to as "The Framework," is vastly different from earlier benchmarks and standards. It provides a sound, evidence-based foundation for standards drawing on current scientific research, including research on the ways students learn science effectively and identifies the science all K-12 students should know. (https://www.nextgenscience.org/framework-k-12-science-education)

Here are the next two statements:

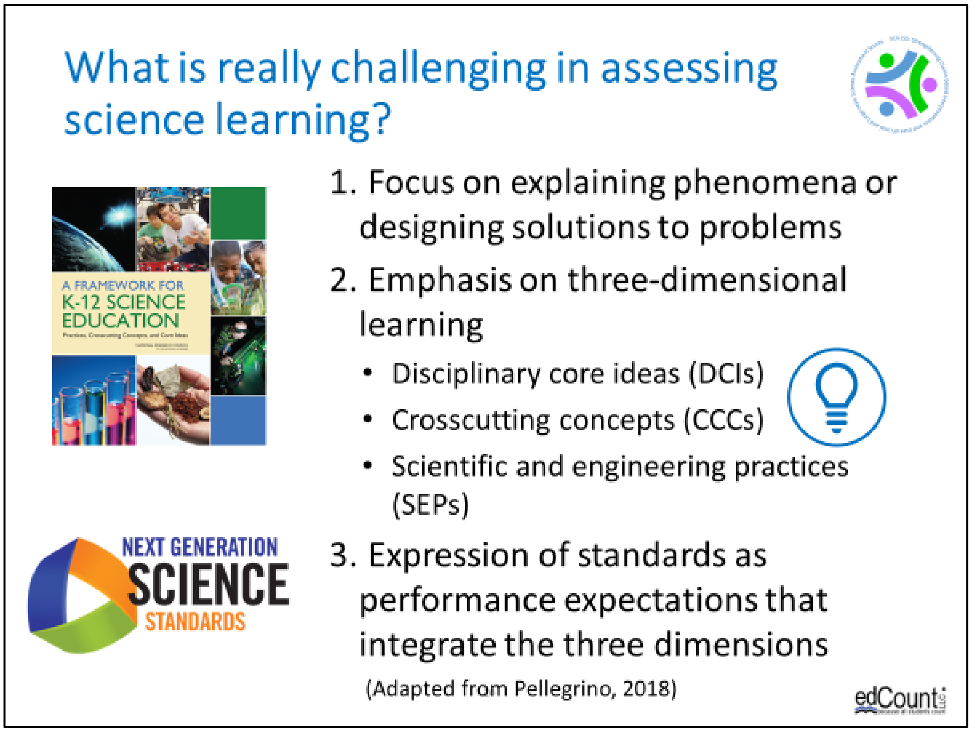
- Designed to be assessed in real-world context. Previous state standards or NGSS?
- Often assessed with multiple choice questions that emphasized definitions. Previous state standards or NGSS?

Assessment in the service of science learning based on the NGSS and in keeping with the vision of The Framework requires much more than twenty multiple choice questions that rely on rote memory. The assessments must utilize meaningful and relevant, real-world context, and we only need to look to the world around us to experience its wonder, phenomena, and design challenges to define those contexts.

Here are the last two statements:

- Engineering integrated with science. Previous state standards or NGSS?
- Engineering often excluded. Previous state standards or NGSS?

You've got it! Integration is key and has the benefit of getting students actively engaged in "doing the science!" Each NGSS performance expectation combines a science or engineering practice, disciplinary core idea, and crosscutting concept into a single statement of what is to be taught and assessed at a grade level or grade band.



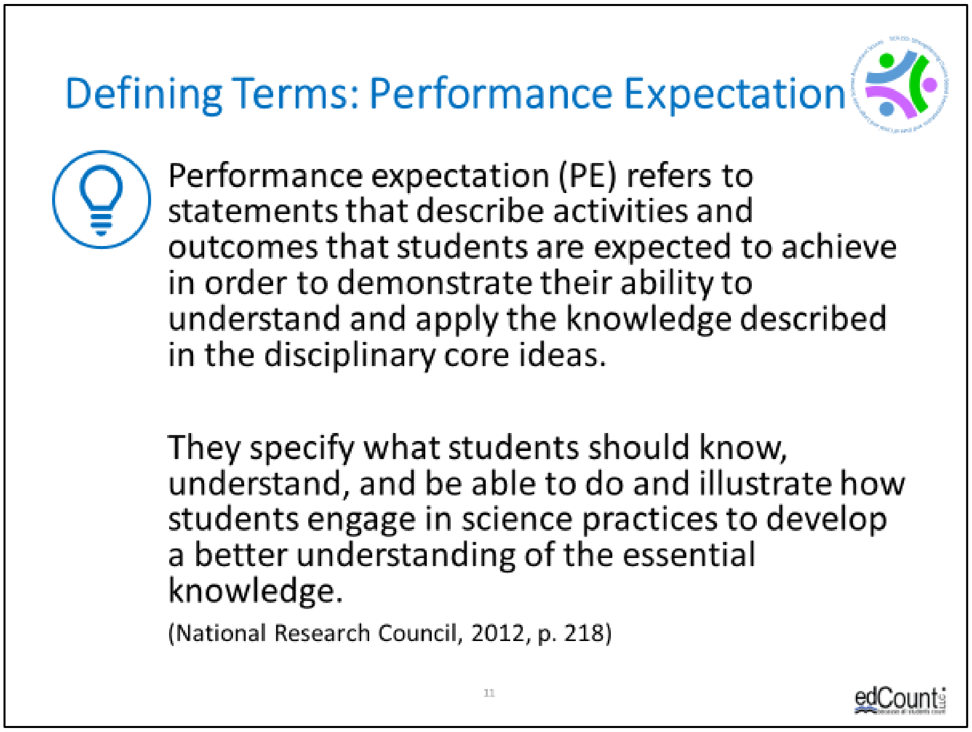
So what is challenging in assessing science learning based on the performance expectations of the NGSS or your state's three-dimensional science standards or indicators? The new vision for K-12 science education calls for engaging students in three-dimensional learning and, thus, requires us to use new ways to assess this learning.

As we have shared, phenomena and design problems play a key role in the NGSS. These have traditionally been a missing piece in science education, which too often has focused on teaching general knowledge or facts that students can have difficulty applying to real-world contexts. Learning to explain phenomena and solve problems are central reasons why students need to engage in the three dimensions of the NGSS during instruction. Now, measuring science understanding is more than measuring content knowledge.

Meaningful measurement of increased and more in-depth understanding of these complex, three-dimensional standards over the course of students' K-12 educational experience requires that assessment tasks, and the evidence of learning they elicit, demonstrate students' ability to use and apply knowledge through the integration of the three dimensions of science understanding.

The tasks must ask students to apply knowledge of disciplinary core ideas and crosscutting concepts while engaging in a science or engineering practice. And students will need multiple and varied assessment opportunities to demonstrate their science learning on specific performance expectations, standards, or indicators for a given grade.

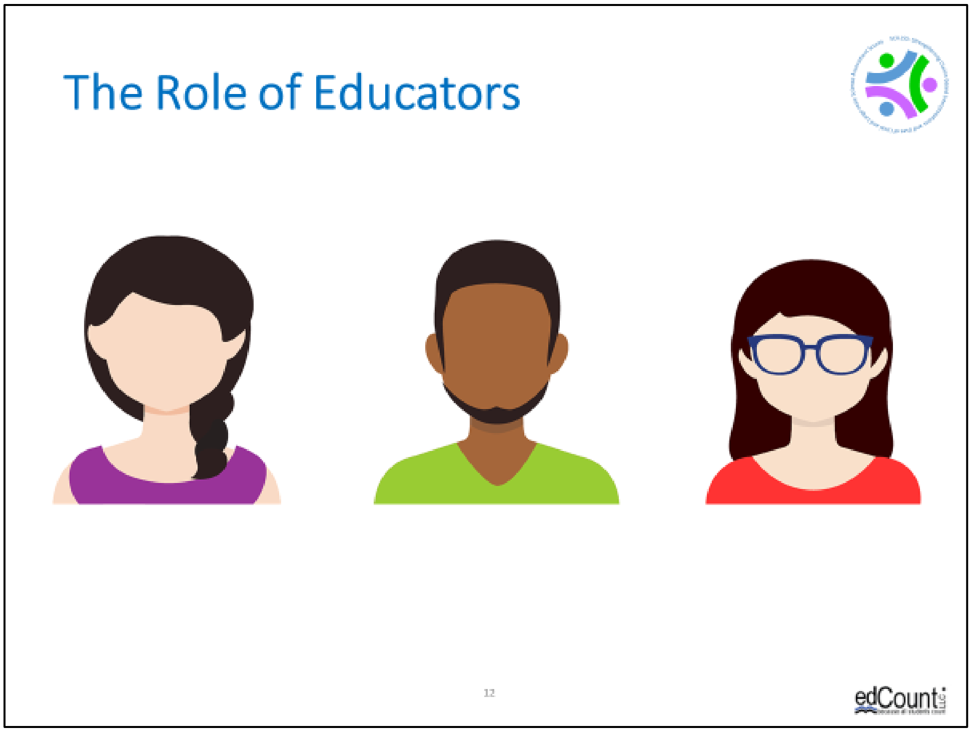
Please refer to the interactive glossary for definitions of Disciplinary Core Ideas, or DCIs, Crosscutting Concepts, or CCCs, and Science and Engineering Practices, or SEPs.



Performance expectation (PE) refers to statements that describe activities and outcomes that students are expected to achieve in order to demonstrate their ability to understand and apply the knowledge described in the disciplinary core ideas.

They specify what students should know, understand, and be able to do and illustrate how students engage in science practices to develop a better understanding of the essential knowledge.

PEs increase in sophistication across the grades. PEs at the higher grades reflect deeper understanding, more highly developed practices, and more complex reasoning (The Framework, p. 228). Thus, with appropriate learning experiences aligned to the PEs, students' conceptual knowledge increases in depth and sophistication as does their use of the practices.



All along every student's educational pathway, beginning in kindergarten and extending through high school, you—educators—are a critical component of an educational system that works to achieve improved student outcomes for every student.

As educators, you impact student learning, bringing your expertise, knowledge, practices, and experience to bear in each and every teaching and learning opportunity you share with your students.

As educators, you facilitate students' progression of science learning by building on their existing knowledge and by being "learner centered."

You understand that you must teach for understanding, rather than retrieval. You see all students as capable of learning and value their ability to transfer that learning to new experiences and contexts.

You are valued by your students, their families and guardians, your fellow colleagues, and your communities.

A rich science education has the potential to capture students' sense of wonder about the world and to spark their desire to continue learning about science throughout their lives.

Research suggests that personal interest, experience, and enthusiasm—critical to children's learning of science at school or in other settings—may also be linked to later educational and career choices (The Framework, 27-30).

There is increasing recognition that the diverse customs and orientations that members of different cultural communities bring both to formal and to informal science learning contexts are assets on which to build—both for the benefit of the student and ultimately of science itself.

For example, researchers have documented that children reared in rural agricultural communities, who experience intense and regular interactions with plants and animals, develop more sophisticated understanding of ecology and biological species than do urban and suburban children of the same age (The Framework, 31-33).

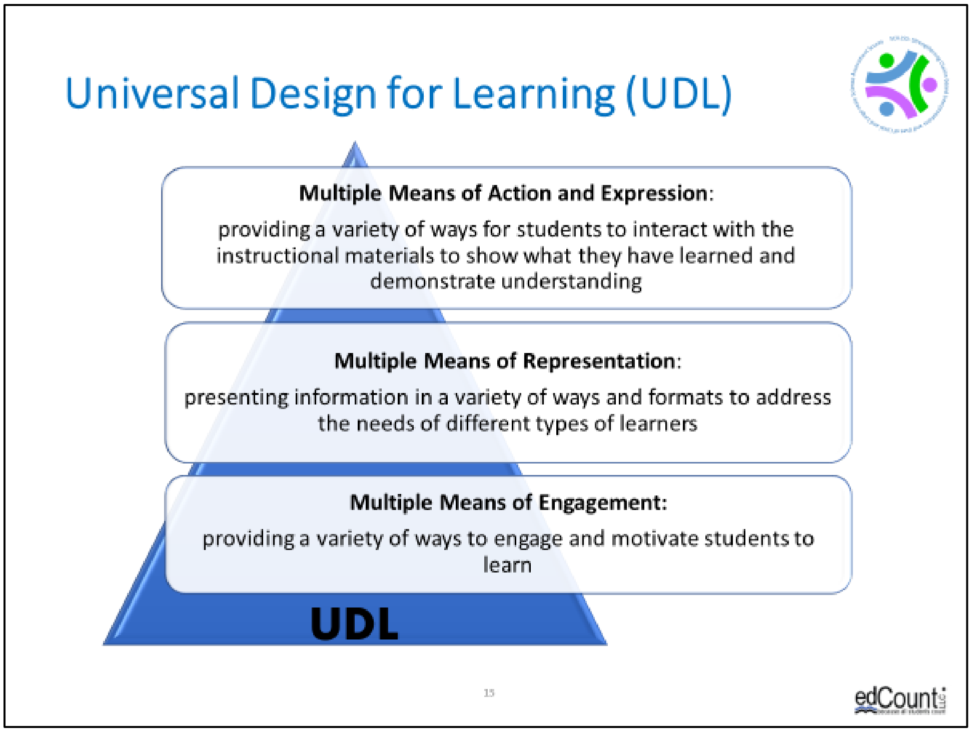
Equity in science education requires that all students are provided with fair and equal opportunities to learn science and become engaged in science and engineering practices; with access to quality space, equipment, and teachers to support and motivate that learning and engagement; and adequate time spent on science.

In addition, the issue of connecting to students' interests and experiences is particularly important for broadening participation in science.

Creating equity in science education can be supported by the exploration of inclusive teaching and learning opportunities for students.

Inclusion is not a place, but rather a systemic approach to uniquely addressing student learning and social engagement within the same instructional frameworks and settings designed for the whole school community.

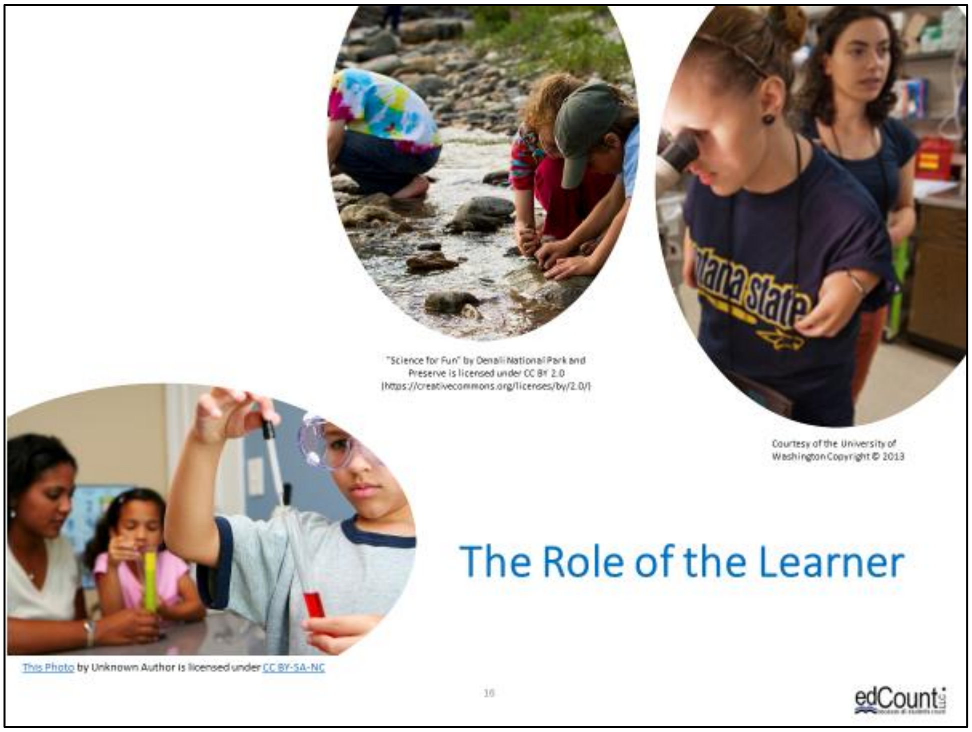
Given that research has delineated benefits for all students when inclusive practices are employed within classrooms, it is our work as educators to investigate and implement educational contexts and strategies, techniques, curricula, and assessment methods that support effective inclusion and benefit all learners (National Council on Disabilities, 2018).



Equitable opportunities in science education are promoted through Universal Design for Learning (UDL). In fact, "universal design" is a phrase borrowed from architecture referring to accessibility for all.

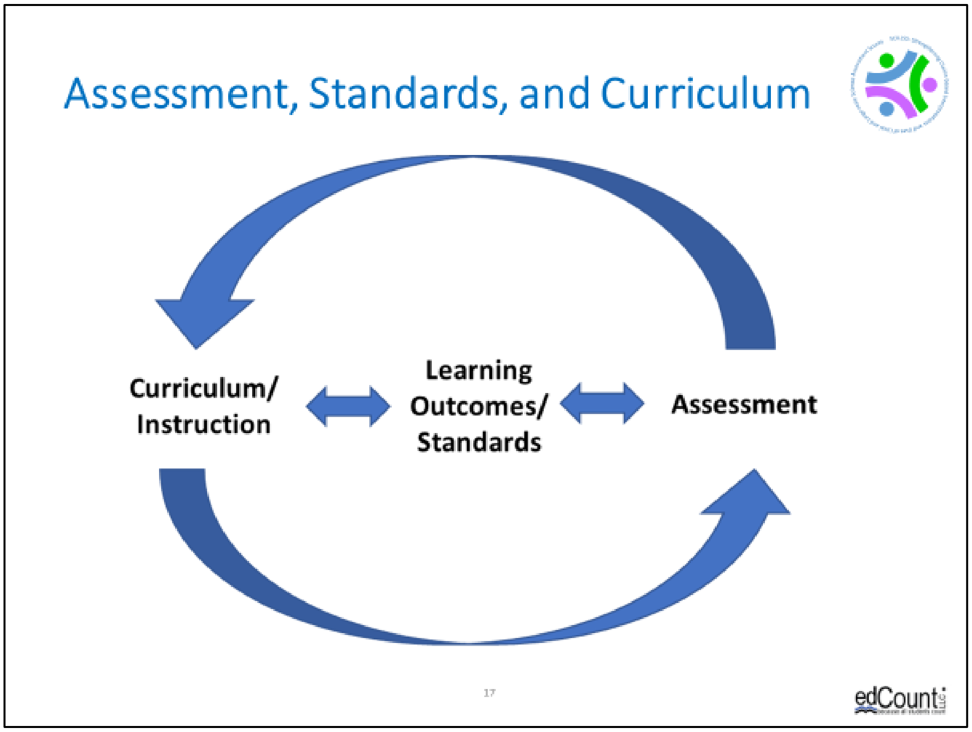
Just as a building must be accessible to all persons regardless of any physical limitations, assessments must be accessible to as many students as possible, regardless of any special needs. Just as wheelchair ramps are built into curbs and doorknobs have been replaced by levers, we need to provide instruction and create assessments that are as free from barriers as possible and allow all students to demonstrate what they know and can do.

The three principles of UDL presented here emphasize the importance of creating a variety of ways to engage and motivate students to learn, providing multiple ways to present information, and allowing students to interact in different ways with materials. Consider how you may incorporate these principles into teaching and learning opportunities and assessments to promote equity and to allow all students to demonstrate what they know and can do (http://udlguidelines.cast.org/).



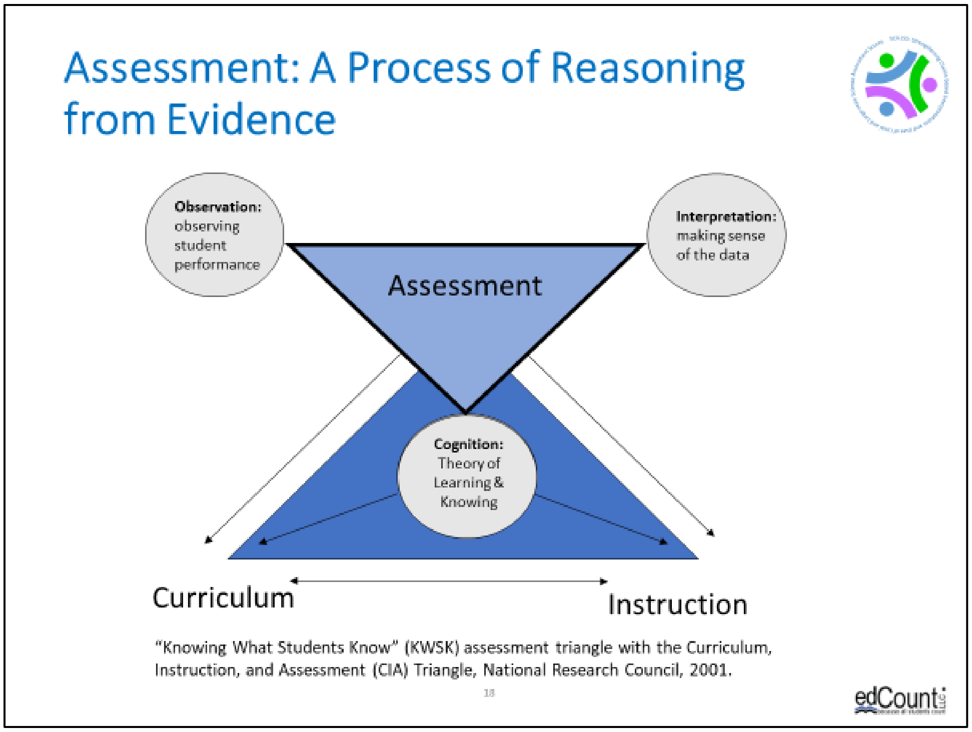
Having addressed the role of the educator, we now turn to the equally essential role of the learner. Learning is a "constructive" process that is active and not passive. Students arrive with prior knowledge, and often with an incomplete understanding of the science content. In addition, students arrive with different interests, identities, and perspectives that contribute to diverse scientific understanding and problem-solving.

To benefit from instruction, learners need to gain an understanding of what they know, what they don't know, and where they may have misconceptions. Knowing what you know and what you don't know is key. Every learner's view of his or her own knowledge and abilities is central to acquiring new and more sophisticated understandings related to science.



A model of learning can serve as a unifying element—a nucleus—that brings coherence to curriculum, instruction, and assessment. This cohesive function is a crucial one because educational assessment does not exist in isolation but must be aligned with curriculum and instruction if it is to support improved learning outcomes. Coherence is at risk if the assessment does not address the topics that are included in the instruction, and consequently, does not provide results that reflect how students are doing in relation to these topics.

At the same time, if students are not doing well on assessments because of material that is not a goal of the curriculum, or if teachers have to spend time focusing on material that is not relevant in order to support their students in doing well on assessments, then these assessments are not helpful to teachers and can, in fact, detract from student learning by taking time away from instruction.



"Knowing What Students Know" (KWSK) assessment triangle with the Curriculum, Instruction, and Assessment (CIA) Triangle, National Research Council, 2001.

The Curriculum, Instruction, and Assessment Triangle rests on cognition, defined as a "theory or set of beliefs about how students represent knowledge and develop competence in a subject domain" (National Research Council, 2001, p. 44).

For the NGSS, the cognition to be assessed consists of the practices, the crosscutting concepts, and the disciplinary core ideas as they are integrated in the performance expectations.

Over the past several decades, powerful insights have been gained into how students represent knowledge and develop competence in specific domains, as well as how tasks and situations can be designed to provide evidence for inferences about what students know and can do across a full range of performance.

Educators must design and implement a coherent system of curriculum, instruction, and assessment that is built from the same model and expectations for teaching and learning.

In pursuing a new form of multi-dimensional science assessment, it is important to remember that assessment is a system composed of three interconnected elements—cognition, observation, and interpretation—and that assessments function within a larger system of curriculum, instruction, and assessment.

The Cognition vertex is a link between assessment, curriculum, and instruction. It represents what educators aim to teach students, and it guides the focus of the curriculum materials and assessments.

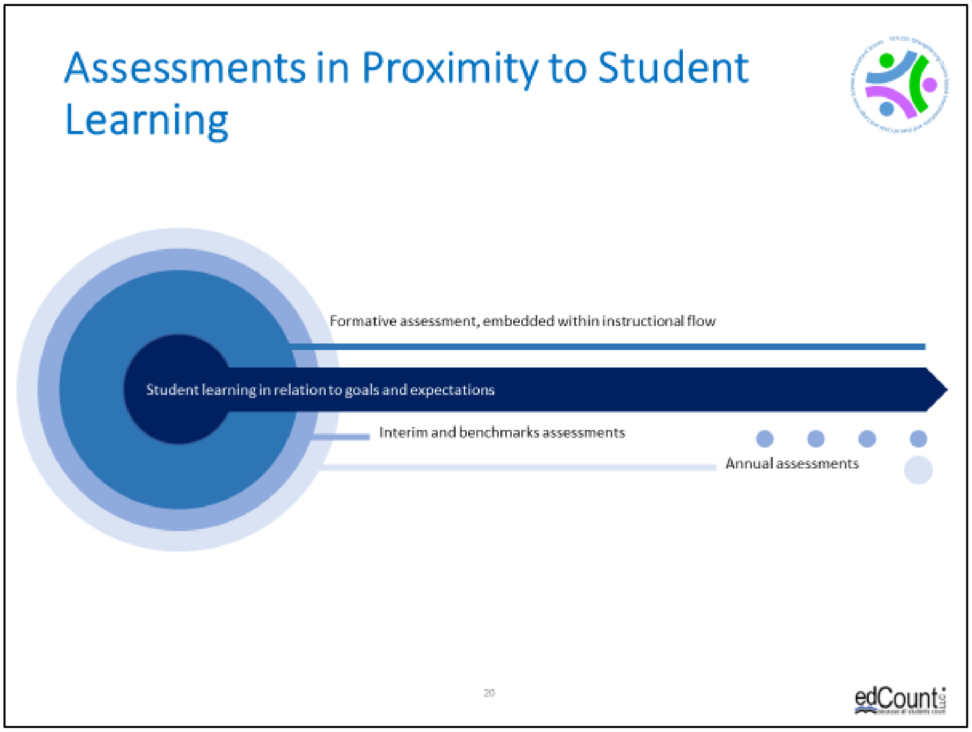
Observations, which include assessment tasks along with the criteria for evaluating students' responses, must be carefully designed to elicit the knowledge and cognitive processes defined in the model of student learning. The design and selection of the tasks need to be tightly linked to the specific inferences about student learning that the assessment is intended to support. For the Observation vertex, educators determine what tests will be developed, how the tests will be administered, and how the tests will be scored.

The third vertex of the triangle is Interpretation, meaning the methods and tools educators use to reason from the observations that have been collected. It focuses on what educators do with the observed evidence, whether it is engaging students in the assessment process, informing next steps for instruction, adapting curriculum, or assigning grades.

Radically changing one of these elements and not the others runs the risk of producing an incoherent system. All of the elements and how they interrelate must be considered together. Thus, designing an assessment is a process in which every decision should be considered in light of each of these three elements.

The Evidence-Centered Design (ECD) approach, developed by Mislevy and colleagues (see, e.g., Almond et al., 2002; Mislevy, 2007; Mislevy et al., 2002; Steinberg et al., 2003), is one framework for developing assessments that takes into account the logic embedded in the assessment triangle and closely follows the evidentiary reasoning logic spelled out by the National Research Council (NRC) assessment triangle.

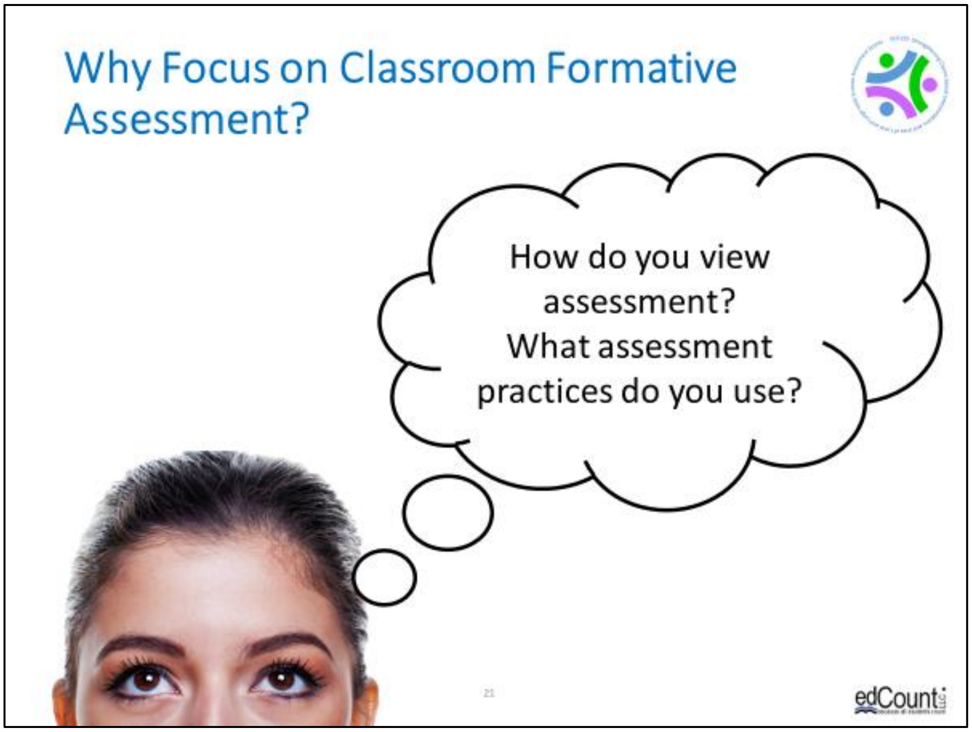
Recall, evidence that tests reflect the concepts that were meant to be measured, is one of the critical sources of information necessary to support valid interpretations of test scores (AERA, APA, and NCME, 2014). Alignment is about coherent connections across various aspects within and across a system (Forte, 2013a, 2013b).



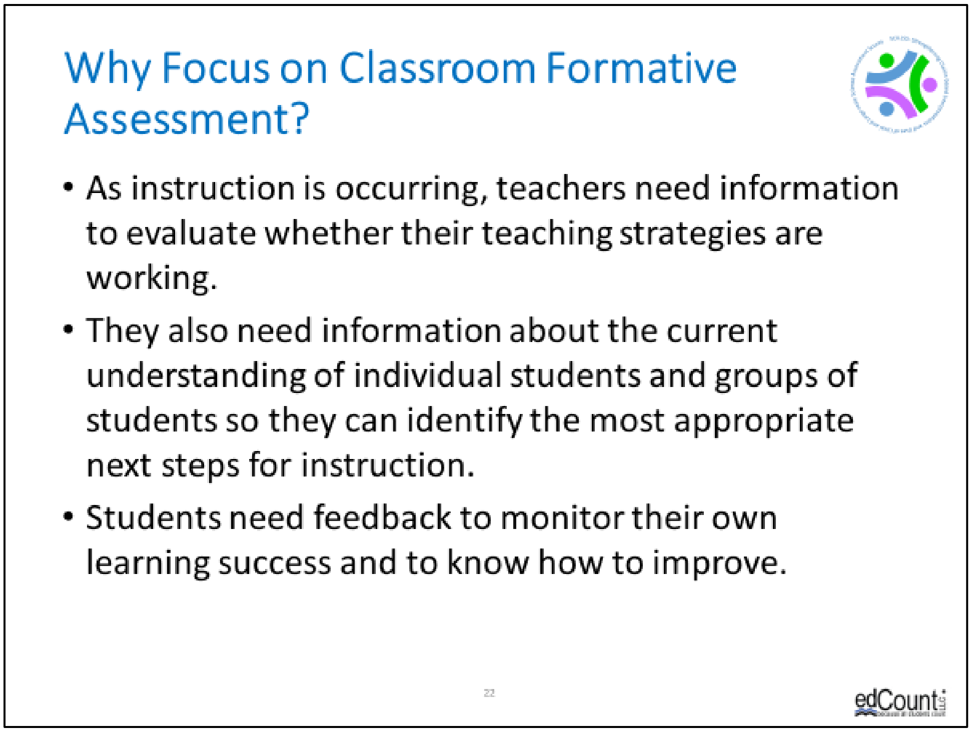
Student learning is the goal; it's at the center of a coherent system of aligned standards, curriculum, instruction, and assessment.

This graphic displays a series of concentric circles. The dark blue center circle is representative of student learning in relation to the goals and expectations. Each additional circle, as shown in the lighter blues moving outward, represents the various types of assessments that are used within the curriculum, instruction, and assessment system.

Most proximal to student learning is formative assessment, which is embedded in the instructional flow and which educators can use to make immediate and ongoing adjustments to their teaching to improve student learning. Next, you have your interim and benchmark assessments and then your annual assessments. These assessments may serve different purposes and uses and are, in some cases, more distant from when learning takes place. The key idea here is that all these assessments in this system are based on common student learning outcomes and are complementary in their measurement of what we expect students to know and be able to do.



Why is a focus on classroom assessment necessary? Think about your own views of assessment, particularly formative assessment in the classroom. How do you view assessment? What assessment practices do you use? Please pause the presentation and reflect on these questions individually or with colleagues.



Classroom assessments can provide information at a finer grain level for individual students by allowing educators to focus on students' current understandings of specific DCIs, CCCs, and SEPs or integration of the dimensions at varying levels. At the class level, classroom formative assessments can help educators identify any patterns or trends that emerge with regard to students' demonstrated knowledge, skills, and abilities or misconceptions related to the standard/performance expectation. Based on these observed patterns or trends, educators can consider how to adjust instruction to address students' learning needs.

The very first step in designing or adopting an assessment is to establish a clear purpose for doing so. Teachers may pose questions to take stock of what students already know at the beginning of a lesson or to determine whether students fully understand a concept or need additional instruction. A district or a state may require students to take specific tests for the purpose of monitoring achievement as part of its accountability system or to provide teachers with information meant to inform classroom instruction. To understand our purposes for assessment, we must consider what it is we want to know, when we need to know it, and what we will do with the information the test provides.

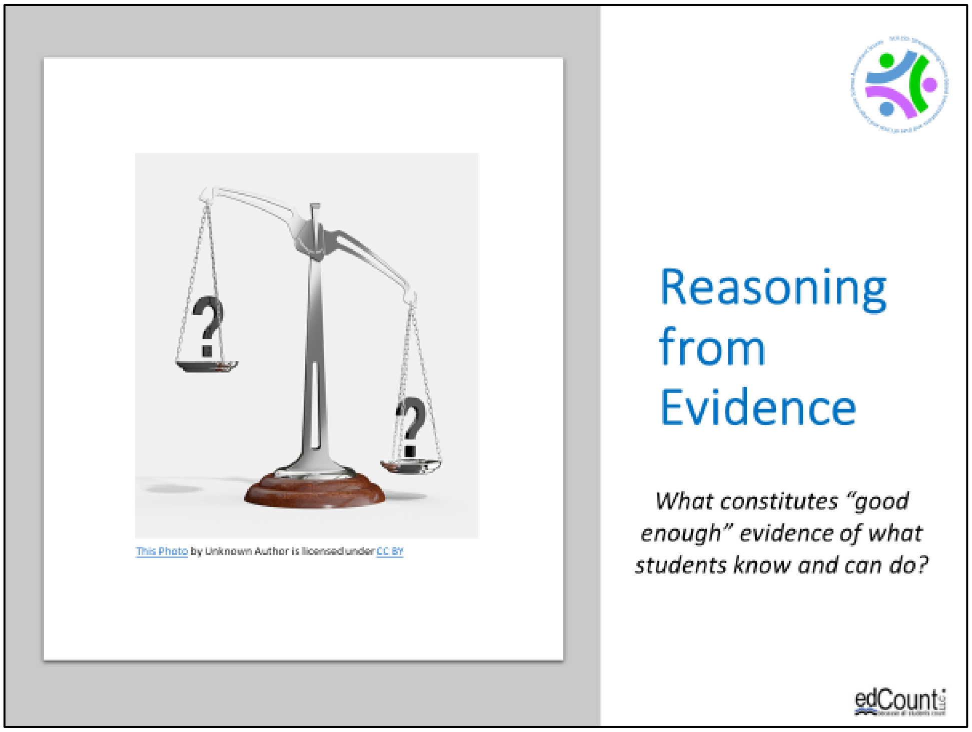
- Is our purpose to gain an understanding of what students are struggling with? Is it to determine if students have learned the material?
- Is it to compare students? Is it to assign a grade?
- Is it to help figure out what the next steps should be in terms of instruction?

Once we understand the purpose of the assessment, we can determine the timing for the assessment.

Should the assessment be administered before instruction? In the middle of instruction or after a series of lessons? Or perhaps after instruction at the end of an instructional unit or quarter?

Another key consideration for intentional assessment design is determining what types of evidence need to be collected. What aspects of students engaging in the task would allow us to say something about the student and how well do these aspects reflect student understanding?

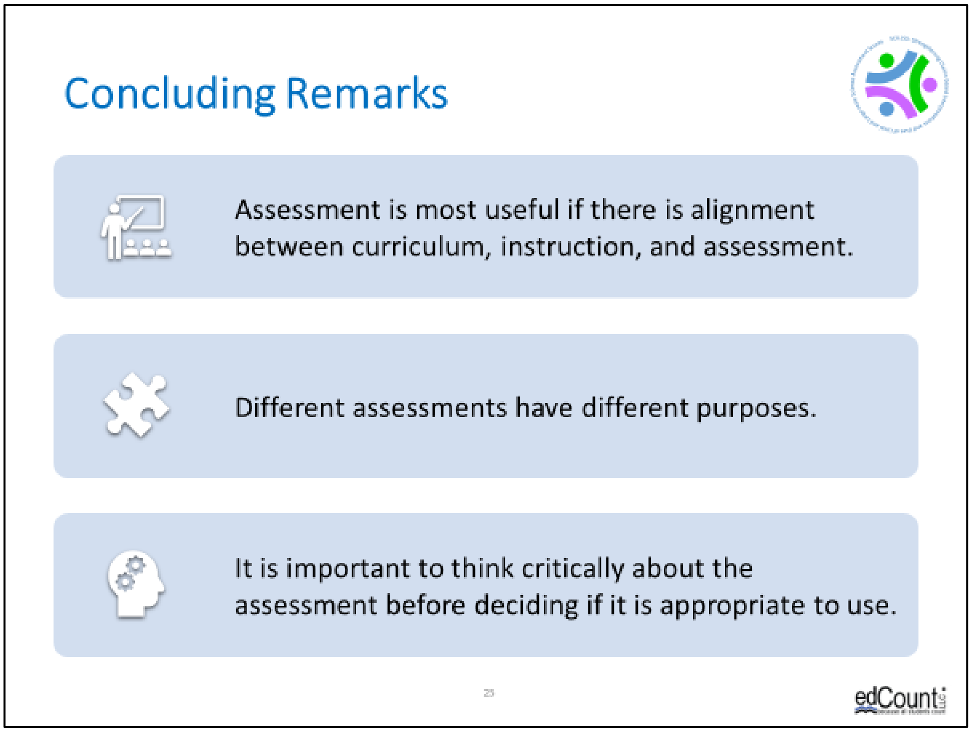
For example, if an educator wants to determine students' ability to design a solution to an everyday problem, what is the best way to gather that evidence? Is it using a series of multiple-choice items? Is it the accuracy of the response? Is it students' ability to create a model or demonstration? Is it their ability to explain and support their ideas? Or is it how quickly they can create a solution? The design of the task, when it is administered, and the types of evidence it provides need to ultimately support the educators' purpose for giving the assessment.



Educators must consider what constitutes "good enough" evidence of what students know and can do, and how to use the evidence to make decisions. What if you wanted to gather information on a student's ability to swim? If all you know about a student is that she is able to identify different swimming strokes, would you consider that student able to swim? Would that be enough proof for you to let her go in water?

Maybe. Knowing that a student can identify swimming strokes might make you more comfortable with allowing her to swim but having observed her swim previously or knowing that she is on the swim team might make you more confident in your decision.

That's the thing about evidence, you might always be gathering it, but to truly know the degree to which students understand a concept, you need to make sure you are gathering the right evidence in the right way for what you want to know.



While in this module we focus on assessment, we have attempted to highlight the critical role that curriculum and instruction play in the development of students' knowledge and skills as well as the interconnections between curriculum, instruction, and assessment. This is especially true for formative assessment as it interacts with curriculum and instruction at the closest level, providing feedback on the acquisition of the learning goals and areas where instructional practices might be improved.

A key takeaway from this module is the understanding that different assessments have different purposes, and it is important to think critically about an assessment before deciding if it is appropriate to use. Educators must consider:

Intentional assessment design requires close attention to the alignment among student learning outcomes, curriculum, instruction, and assessment and the purposes that define why we assess. Always ask yourself, is the purpose of the assessment task appropriate for the situation? What will I learn about my students from using this task and how does this align to what I have been teaching? Is the evidence enough to support the claims about the student? And, are there any other explanations for why a student might not do well on this task?

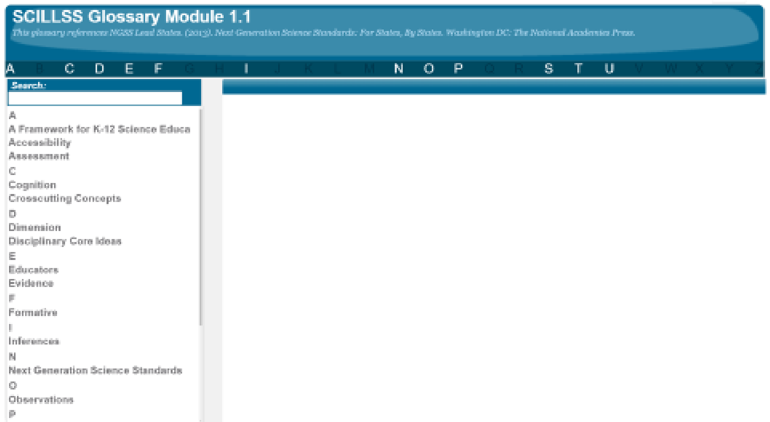
In Module 1.2, we will explore how to design three-dimensional classroom science tasks to address the vision of The Framework and the NGSS and to support educators in shifting their curriculum, instruction, and assessment practices to improve student learning outcomes in science.

Finally, we offer additional resources that may be helpful to anyone interested in learning more about the concepts presented in this module. A glossary of terms and our reference list follow.

Thank you for your engagement in this first chapter of the SCILLSS digital workbook on designing high-quality three-dimensional science assessment tasks for classroom use.

SCILLSS Glossary

Please refer to the SCILLSS Glossary for operational definitions of terms used.



Web Links

In the web links pod, you can find the following resources:

- "THE AMAZING TRIPLE SPIRAL (15,000 DOMINOES)"
- A Framework for K-12 Science Education
- Next Generation Science Standards
- Universal Design for Learning Guidelines

28

edCount LLC

References

Almond, R. G., Steinberg, L. S., & Mislevy, R. J. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *Journal of Technology, Learning, and Assessment, 1*(5). Retrieved from http://www.jtla.org

American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing*. Washington DC: AERA.

Forte, E. (2013a). *Re-conceptualizing alignment in the evidence-centered design context*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Forte, E. (2013b). *Evaluating alignment for assessments developed using evidence-centered design.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.

Mislevy, R. J. (2007). Validity by Design. *Educational Researcher*, *36*(8), 463 469. https://doi.org/10.3102/0013189X07311660.

Mislevy, R.J., Steinberg, L.S., & Almond, R.G. (2002). On the structure of educational assessments. Measurement: *Interdisciplinary Research and Perspectives, 1*, 3-67.

29

edCount

References

National Council on Disabilities. (2018). (IDEA Series). *The Segregation of Students with Disabilities*. Washington, DC: Author.

National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: The National Academies Press. https://doi.org/10.17226/10019.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13165.

National Science Teaching Association. (n.d.). *Compare NGSS to Existing State Standards*. Retrieved from https://ngss.nsta.org/compare-the-standards.aspx.

NGSS Lead States. (2013). Next Generation Science Standards: For States, By States. Washington DC: The National Academies Press.

Pellegrino, J. W. (2018). *Assessing Science Learning: Some of the Complexities & Challenges* [PowerPoint Slides].

Steinberg, L.S., Mislevy, R.J., Almond, R.G., Baird, A.B., Cahallan, C., DiBello, L.V., Senturk, D., Yan, D., Chernick, H., & Kindfield, A.C.H. (2003). *Introduction to the Biomass project: An illustration of evidence-centered assessment design and delivery capability*. CSE Technical Report. Los Angeles: UCLA Center for the Study of Evaluation.

edCount