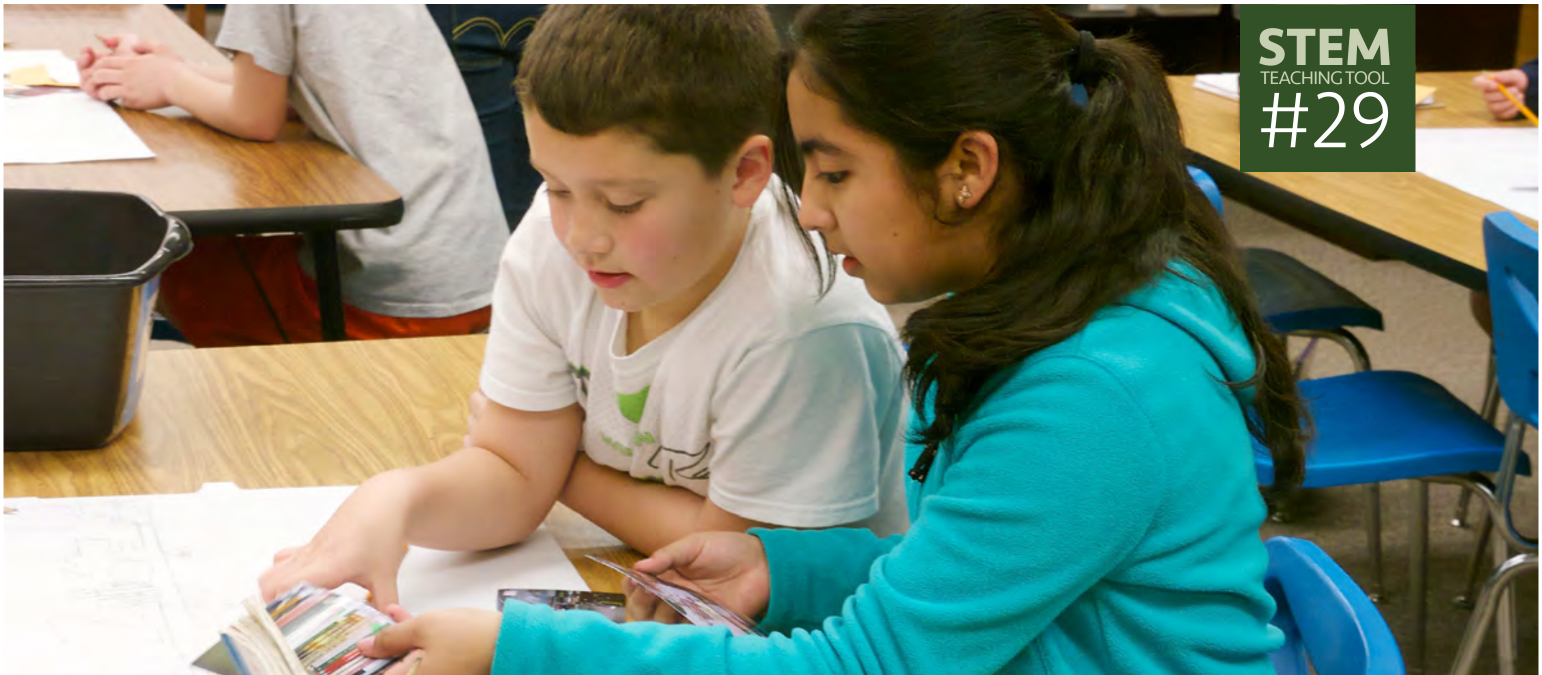

Steps to Designing Three-Dimensional Assessments that Connect to Students' Interests, Experiences, and Identities

This nine-step process is designed to help teams develop three-dimensional assessment tasks in science that connect to students' interests, experiences, and identities. The process, if followed as outlined here, can be expected to take between 4 and 6 hours for a small team to develop a single extended task to be used as part of a unit test. Ideally, teams should begin the process with a basic understanding of the vision of *A Framework for K-12 Science Education* (National Research Council, 2012). However, the process can also be used to help orient educators to the Framework vision.

Three-dimensional assessment tasks allow you to make inferences about how students use their understanding of disciplinary core ideas, science and engineering practices, and crosscutting concepts together to explain phenomena and solve problems. This process includes steps to ensure that your tasks are accessible and engaging to learners from non-dominant communities and to learners with identified learning differences.

Assessing three-dimensional standards means assessing more than just the "process" of science; it means assessing students' proficiency through integrated use of all three dimensions to explain phenomena and solve design challenges. The guidance included here is based on the conclusions and recommendations included in the National Research Council (2014) report, Developing Assessments for the Next Generation Science Standards. In addition, the process incorporates use of the Task Screener developed by Achieve, Inc., to analyze existing assessments.

This document was updated in 2020 and replaces the prior version of this resource.

Funded by the National Science Foundation (NSF). Opinions expressed are not those of any funding agency.

BY WILLIAM R. PENUEL, KATIE VAN HORNE, AND PHILIP BELL REVISED MARCH 2020 (FROM MARCH 2016 VERSION)

Overview of the Steps Outlined in this Document

Throughout this document, we make use of three key terms:

Task: A single, multi-component activity designed to elicit understanding of a performance expectation (or part of one).

Scenarios: The contextual framing of a task in which a phenomenon to be explained or a problem to be solved is presented to students. The phenomenon or problem presented is the focus of the entire assessment.

Prompts: Individual components or questions, all of which are linked to the scenario.

Assessment design is best conducted in teams, where colleagues can work together to clarify learning goals to be assessed, brainstorm phenomena and design challenges that will require application of the disciplinary core ideas of a performance expectation, and review one another's prompts.

This tool is intended to be used in conjunction with other STEM Teaching Tools:

ACESSE Resource D: How to Craft 3D Classroom Science Assessments

ACESSE Resource E: Selecting an Anchoring Phenomenon for Equitable 3D Teaching

STEM Teaching Tool 30: Integrating Science Practices into Assessment Tasks

STEM Teaching Tool 41: Prompts for Integrating Crosscutting Concepts into Assessment and Instruction

If you have never developed a three-dimensional assessment task as a team, begin with STEP 0, "Becoming Familiar with Three-Dimensional Tasks" or engage with ACESSE Resource A and ACESSE Resource B. If you have developed three-dimensional assessment tasks as a team and are familiar with using the Achieve, Inc., task screener, skip to STEP 1.

STEP 0: Become familiar with qualities of three-dimensional assessment tasks

For people just beginning to develop three-dimensional assessment tasks, it is useful to begin with an analysis of existing tasks. One task designed using the process described here is the Swallows Task. It assesses students' understanding of a high school life science performance expectation. As a team, take a minute to read the annotated assessment. Then discuss what you notice about it, especially how it is different from assessments you may have used in the past.

Next, look at the summary of strengths and opportunities, which shows what others who have experience developing three-dimensional assessments have written about it. How are their ideas similar or different from your own? What ideas do you now have about what makes for a good three-dimensional assessment?

You can record your ideas about what makes for a good three-dimensional assessment task and keep them nearby as you go through the rest of the steps in the process.

Achieve, Inc's Science Task PreScreen: Basic Criteria for Three-Dimensional Assessment Tasks

- Is there a phenomenon or problem driving the task?
- Can the majority of the task be answered without using information provided by the task scenario? (answer should be "no")
- Can significant portions of the task be answered successfully by using rote knowledge (e.g., definitions, prescriptive or memorized procedure)? (answer should be "no")
- Does the majority of the task require students to use reasoning to successfully complete the task?
- Does the task require students to use some understanding of disciplinary core ideas to successfully complete the task?
- Do students have to use at least one science and engineering practice to successfully complete the task?
- Are the dimensions assessed separately in the majority of the task?
- Is the task coherent and comprehensible from the student perspective?

STEP 1. Define what you will assess by analyzing relevant sections of *A Framework for K-12 Science Education* and crafting learning claims

Assessment begins with defining what you want to be able to say about what your students know and can do. The *Framework* provides a starting place for defining the understandings that should be assessed at each grade band. The practices chapter (Chapter 3) and crosscutting concepts chapter (Chapter 4) highlight grade 12 endpoints and what is known about progressions across K-12. The disciplinary core ideas chapters (Chapters 5-8) include descriptions for what students are expected to know and be able to do by the end of grades 2, 5, 8, and 12.

Use the *Framework* text to define a set of "learning claims" that you want to be able to make about what students know and can do. A claim is more than just a phrase that references a concept (e.g., "ecosystem stability"). For example, "A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions," is a statement derived from the 12th grade expectation for LS2C. This particular claim is not sufficient to develop a learning claim for an assessment, as it does not account for performance related to the other two dimensions (Science and Engineering Practices and Crosscutting Concepts), as shown in the example below. The performance expectations can be the basis for developing claims. However, many assessment tasks will only assess part of a performance expectation. We may choose parts that fit together, for example, because they are necessary to explain a particular phenomenon, or because conceptually they are related.

Below, we present a performance expectation in the form of a claim, and show the relevant Framework text that makes up the claim. The text in this table is the same text that is in the "connections" boxes of the NGSS.

Claim	Relevant Framework Text
Students can evaluate the claims, evidence, and reasoning that the complex interactions in ecosystems maintain relatively consistent numbers and types of organisms in stable conditions, but changing conditions may result in a new ecosystem. Associated PE: HS-LS2-6	A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions. If a modest biological or physical disturbance to an ecosystem occurs, it may return to its more or less original status (i.e., the ecosystem is resilient), as opposed to becoming a very different ecosystem. Extreme fluctuations in conditions or the size of any population, however, can challenge the functioning of ecosystems in terms of resources and habitat availability. (LS2C: Ecosystem Dynamics, Functioning, and Resilience) Evaluate the claims, evidence, and reasoning behind currently accepted explanations or solutions to determine the merits of arguments. (SEP: Engaging in Argument from Evidence) Much of science deals with constructing explanations of how things change and how they remain stable. (CCC: Stability and Change).

STEP 2: Analyze the facets of the claim to be assessed

Developing a deeper understanding of the components of the performance expectation is needed, both to select a phenomenon or design challenge for the assessment and to develop a sense of what is important to score in an assessment.

An analysis of the facets of the performance expectation is the second step in assessment design. A facet is a small observable piece of knowledge or a strategy that a student uses to make sense of a problem (Minstrell, 1992). Facets can be about any dimension of science learning, and can be written as claims, such as "Students can define the boundaries of ecosystems on the basis of direct and indirect interactions among organisms," or "Argument from evidence involves evaluation of given claims."

The NGSS Evidence Statements are one representation of the facets of a performance expectation. However, actively analyzing facets—including the underlying conceptual model elements that are desired—can be a valuable way for a team to develop a shared understanding of the learning targets for an assessment task.

Making "sticky notes" with the facets on them can help teams keep track of whether the assessment prompts you have designed elicit each facet.

Sample Facets (DCI Only)	Relevant Framework Text
Ecosystems interactions can create stability, in terms of the numbers and types of organisms. The numbers are "relatively" constant, meaning they do bounce up and down within a range and, in some cases, seasonally. The constancy can be over a long period, provided that the conditions of an ecosystem are stable. • One "condition" might be the absence of major disturbances to the ecosystem.	A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions. If a modest biological or physical disturbance to an ecosystem occurs, it may return to its more or less original status (i.e., the ecosystem is resilient), as opposed to becoming a very different ecosystem. Extreme fluctuations in conditions or the size of any population, however, can challenge the functioning of ecosystems in terms of resources and habitat availability. (LS2C: Ecosystem Dynamics, Functioning, and Resilience)
We can conclude an ecosystem is resilient if something disturbs it, and it returns to its original status. • Ecosystems, when perturbed, will find a new equilibrium that may resemble a previous state. Ecosystems have characteristics/factors that can be examined as indicators of change. • We can observe and measure the effect of "modest" disturbances. • "Original status" implies use of an agreed-upon range of factors that count as typical for a given ecosystem. "More or less" implies ecosystems will be affected by disturbances to varying degrees. Here, "status" likely relates to the richness and abundance of different species in the ecosystem and the dynamics of the ecosystem.	A complex set of interactions within an ecosystem can keep its numbers and types of organisms relatively constant over long periods of time under stable conditions. If a modest biological or physical disturbance to an ecosystem occurs, it may return to its more or less original status (i.e., the ecosystem is resilient), as opposed to becoming a very different ecosystem. Extreme fluctuations in conditions or the size of any population, however, can challenge the functioning of ecosystems in terms of resources and habitat availability. (LS2C: Ecosystem Dynamics, Functioning, and Resilience) Evaluate the claims, evidence, and reasoning behind currently accepted explanations or solutions to determine the merits of arguments. (SEP: Engaging in Argument from Evidence) Much of science deals with constructing explanations of how things change and how they remain stable. (CCC: Stability and Change)

STEP 3: Choose a phenomenon or a design challenge for the task—framed through a compelling scenario

Every three-dimensional assessment requires a scenario, in which students are presented with a phenomenon or a design challenge. Phenomena require the use of science practices to explain, and design challenges require the use of engineering practices to solve. Both should require students to apply their understanding of targeted disciplinary core ideas. This particular requirement means that developing a task will involve time looking for a good phenomenon or design challenge. The task's scenario should be engaging, relevant, and accessible to a wide range of students (Achieve, 2018). (See also ACESSE Resource D and STEM Teaching Tool #28 and #42.)

ACESSE Resource E provides a process for selecting a phenomenon that can anchor either a three-dimensional assessment or sequence of lessons. That process entails:

- Analyzing the focal performance expectation(s) (STEP 2 above)
- Brainstorming candidate phenomena — and consider different kinds of phenomena for specific equity purposes
- Drafting a student explanation of 6-7 candidate phenomena to choose from
- Administering a student survey to assess students' interest in the phenomenon
- Choosing a phenomenon and developing an explanation of the phenomenon you expect students to be able to give at the end of a sequence of instruction (STEP 4 below)

A parallel process could be used to select an engineering design challenge. If phenomena and design challenges are a new idea to you and your team, ACESSE Resource E includes a game you can play together, "The Phenomenon Game."

Choosing a phenomenon that is of interest to students and that is personally relevant to them, to help ensure that students will engage with and persevere through completion of a task. Choosing a phenomenon where scientists or engineers from students' own communities are involved can help students from non-dominant backgrounds identify with science and engineering. Therefore, gathering and analyzing evidence of student interest is a strategy for promoting equity through the choice of assessments. The example below is based on a phenomenon chosen through the process described above. Other examples of scenarios can be found here. This particular phenomenon, though not local to students, was of particular interest to girls and to African American and Latinx students. In addition, it requires the use of facets of DCIs to explain. The complete assessment can be found here.

Phenomenon and Associated Facets	Scenario
Phenomenon: *In the Serengeti Plain, the wildebeest plays a special role in the ecosystem. Its numbers determine the numbers of many other organisms in the ecosystem. Changes to its numbers over time have greatly affected the abundance and diversity of the plain.* *Conservation efforts focus on protecting the wildebeest, because of its special role in the ecosystem. At the same time, the wildebeest threaten livestock that local people need to feed their families.* Explaining this phenomenon and solving its associated design challenge requires students to apply their understanding about how interactions create stability over varying periods of time, as well as how changes actually have a dramatic effect on organisms, including through indirect interactions. The phenomenon also affords students the opportunity to examine factors that are indicators of change.	**Phenomenon-Driven Scenario** A keystone species is a plant or animal that plays an important and unique role in how an ecosystem functions. Without the keystone species, the ecosystem would be very different. One way scientists identify keystone species is to look at how changes to their abundance (number) affect other organisms. Often , when the number of a keystone species changes, there are many indirect effects and of changes to the ecosystems. Answer the questions below to evaluate the claim that the wildebeest in the Serengeti is a keystone species. Draw on evidence you have gathered about the Serengeti to support your evaluation of the claim. **Design Challenge-Driven Scenario** Some conservationists argue that the wildebeests should be protected on the Serengeti, even though the size of the wildebeest population is large, in order to protect the ecosystem in the future. Some local governments have sought to get rid of wildebeests in their countries, because the animals threaten livestock (e.g., cows) that people depend on to eat.

STEP 4: Write a complete student explanation of the phenomenon or solution to the design challenge

Before developing specific prompts for the task, it is useful to write a complete student explanation of the phenomenon or solution to the design challenge. The explanation or solution should be what you expect students to write. They are the "answers" to the questions or prompts that you will develop in your assessment. Write them as an emerging bilingual or multilingual student who had mastered the performance expectation might write an explanation.

Scenario	Complete Student Explanation
A keystone species is a plant or animal that plays an important and unique role in how an ecosystem functions. Without the keystone species, the ecosystem would be very different. One way scientists identify keystone species is to look at how changes to their abundance (number) affects other organisms. Often when the number of a keystone species changes, there are many indirect effects and changes to the ecosystems. Answer the questions below to evaluate the claim that the wildebeest in the Serengeti is a keystone species. Draw on evidence you have gathered about the Serengeti to support your evaluation of the claim.	The big growth of wildebeest population after the elimination of Rinderpest led to big direct and indirect changes on the ecosystem. More wildebeest meant there was more food for their predators, such as lions and hyenas. The larger numbers of wildebeest meant there was less grass. They reduced the average height of grass in the plains by several inches. The lower grass levels meant less fuel for fires that periodically burned the plains. With fewer fires, the seeds of trees were allowed to grow, and eventually the plain became more covered with trees. This in turn led to a larger giraffe population. So, these direct and indirect effects led to big changes in the ecosystem, all because of the numbers of wildebeest. Some evidence points to the fact that elephants are also keystone species in the Serengeti.

STEP 5: Use the Science and Engineering Practices and Crosscutting Concepts Tools to develop individual prompts

A key challenge for many teams is to develop tasks that elicit students' grasp of the science and engineering practices and their understanding of crosscutting concepts. Two tools exist to help develop prompts (individual questions) that do so:

STEM Teaching Tool 30: Integrating Science Practices into Assessment Tasks
STEM Teaching Tool 41: Prompts for Integrating Crosscutting Concepts into Assessment and Instruction

You can use the first tool to organize a sequence of prompts that make sense for students, and then adapt prompts from the second tool in appropriate places so that students can show their understanding of crosscutting concepts.

There are multiple task formats for the science and engineering practices. Having multiple task formats to choose from allows for variety in assessment prompts. The formats vary in how challenging they are likely to be for students, too. Some require students to construct knowledge with very little support from the prompt. Others could be used to build questions where students select from available responses (multiple choice). Most assessments will rely on more than one practice, because multiple practices are almost always needed to make sense of phenomena and solve problems (see STEM Teaching Tool #3).

As you develop prompts, check to make sure that all of the elements of the complete explanations are elicited. In the example below, the evidence presented comes from lessons taught over the course of 10 days, as part of a longer-term investigation of the Serengeti ecosystem. In most stand-alone assessments, the data would be provided to students as part of the scenario, as in the Swallows Task.

Scenario (Phenomenon)	Prompts
A keystone species is a plant or animal that plays an important and unique role in how an ecosystem functions. Without the keystone species, the ecosystem would be very different. One way scientists identify keystone species is to look at how changes to their abundance (number) affect other organisms. Often , when the number of a keystone species changes, there are many indirect effects and changes to the ecosystems. Use the space below to evaluate the claim that the wildebeest in the Serengeti is a keystone species. Draw on evidence you have gathered about the Serengeti to support your evaluation of the claim.	**Engaging in Argument From Evidence (Task Formats 6a and 6b from STEM Teaching Tool #30):** What evidence collected throughout the unit supports the claim that the wildebeest is a keystone species? What evidence collected throughout the unit does not support this claim? Stability and Change: Would protecting the wildebeest help to protect the stability of the Serengeti ecosystem? Why or why not?

STEP 6: Integrate questions to assess student interest and identification with science and engineering presented in the scenario

Interest and identification with science can be gauged in the context of an assessment. Gathering evidence of students' perceptions of the personal or community relevance of a scenario can help you monitor equity goals. Assessments can help build relationships with students and help them see themselves in the science, even though they are not typically used for this purpose.

Construct	Example Question (Ideal answers are underlined)
Relevance to Community	What we did in class today matters to people in my community because: (select the option that best describes your feelings) **A. This material is important and people should know about it** **B. This material could improve the lives of people in my community** C. What we did today doesn't matter to people in my community D. Other: (Please write in an answer)
Identification with science	While completing this assessment, I felt (circle all that apply): A. Excited B. Bored C. Frustrated **D. Like a scientist** E. Afraid F. Angry G. Happy People like me do science: (select the option that best describes your feelings) **A. Strongly Agree** B. Agree C. Unsure D. Disagree E. Strongly Disagree I want to learn more science in the future: (select the option that best describes your feelings) **A. Strongly Agree** B. Agree C. Unsure D. Disagree E. Strongly Disagree

STEP 7: Develop ideal student answers and a scoring guide or assessment rubric

Next, develop an answer key. Assign points to facets included in student answers that are linked to the facets identified when analyzing the performance expectation. When developing a scoring guide, make sure that students get the most points for answers that reflect the facets you analyze.

A total score for the assessment can be given as the sum of the individual points. The higher the number of points, the stronger the evidence of mastery of the performance expectation.

Feedback to students can focus on the facets that students still need to develop to construct a satisfactory explanation of the phenomenon or solution to the problem.

Prompt	Student Answers and Scoring Guide
What evidence collected throughout the unit supports the claim that the wildebeest is a keystone species?	**Answers:** The big growth of wildebeest population after the elimination of Rinderpest led to big direct and indirect changes on the ecosystem. More wildebeest meant there was more food for their predators, such as lions and hyenas. The larger numbers of wildebeest meant there was less grass. They reduced the average height of grass in the plains by several inches. The lower grass levels meant less fuel for fires that periodically burned the plains. With fewer fires, the seeds of trees were allowed to grow, and eventually the plain became more covered with trees. This in turn led to a larger giraffe population. So, these direct and indirect effects led to big changes in the ecosystem, all because of the numbers. This makes them a keystone species, because this is the kind of role that keystone species play in an ecosystem. +1 point for each *direct* change to abundance of another organism listed associated with the wildebeest's increase +1 point for each *indirect* change to abundance of organisms associated with the increase of wildebeest +1 point for conclusion that connects evidence to definition of keystone species Note: In this prompt, the majority of points focus on the DCI, but across a task, it is important to award points for performance of each of the dimensions.

STEP 8: Review your task with peers for intelligibility, alignment, and accessibility

Sharing your initial tasks with a colleague and asking them for constructive feedback on how to improve them is a good way to begin. There are many roles your colleagues could play in helping improve your task. They can improve the intelligibility of prompts, so that more students will understand what you are asking. They can help ensure alignment to the targeted performance expectation. And, they can help ensure the accessibility of the task.

Tool or Practice	How It Can Improve Your Task
Have a colleague complete your assessment	Can help ensure you are asking what you think you are asking and getting responses you hope to get from students
Assess whether the scoring guide is aligned with the Evidence Statements for the targeted performance expectations	Can help ensure you are awarding points for what really matters, in terms of what students know and should be able to do
Review the assessment using the Achieve's Task Screener for equity	Can help ensure that the task is accessible and engaging to a wide variety of users

STEP 9: Pilot and revise your assessment

Assessment design requires many cycles of developing, testing, and revising tasks to ensure that you are getting an accurate picture of what students know and can do. It is helpful to pilot test assessments with a small number of students to feed that actual student response data into the revision process.

Carefully review alignment of the different elements presented here (claims, scenarios, application of task formats, and hypothetical/actual student answers) and pilot tasks with students as part of classroom instruction to reveal ways to improve tasks.

One key is to be ready to revise your initial tasks, even when you've put a lot of work into them. Often, the challenge is not with our students, but with the questions that we ask. It is difficult to develop tasks that allow all students to show what they know and can do. Yet it is imperative to do so in order to create fair, valid assessments of students' three-dimensional science proficiency. Also, once you test tasks with students, your hypothetical student responses can be replaced with actual student responses, along with ideas for how to address problematic aspects of student responses.

References

Achieve, Inc. (2018). *Assessment task prescreen*. Washington, DC: Author.

Achieve, Inc. (2018). *Science task screener*. Washington, DC: Author.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. In F. Duit, F. Goldberg, & H. Niedderer (Eds.), *Research in physics learning: Theoretical issues and empirical studies* (pp. 110-128). Kiel, Germany: IPN.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

National Research Council. (2014). *Developing assessments for the Next Generation Science Standards*. Washington, DC: National Academies Press.

This work was created and refined as part of the Advancing Coherent and Equitable Systems of Science Education (ACESSE) and the Research + Practice Collaboratory projects. ACESSE brings together partners from educational research and practice to improve equity by building coherence in science education. The Research + Practice Collaboratory brings educators and researchers together to develop more equitable innovations for STEM teaching and learning. Learn more at cosss.org/ACESSE and researchandpractice.org.

This material is based in part upon work supported by the National Science Foundation under Grant Number EHR-1238253 and by the Gordon and Betty Moore Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funders.