Designing and Developing NGSS-Aligned Formative Assessment Tasks to Promote Equity

Nonye Alozie[a], Phyllis Haugabook Pennock[b], Krystal Madden[c], Sania Zaidi[c], Christopher J. Harris[a], and Joseph S. Krajcik[b]

[a] SRI International
[b] Michigan State University
[c] University of Illinois, Chicago

Alozie, N., Haugabook Pennock, P., Madden, K., Zaidi, S., Harris, C. J., & Krajcik, J. S. (2018, March) *Designing and Developing NGSS-Aligned Formative Assessment Tasks to Promote Equity*. Paper presented at the annual conference of National Association for Research in Science Teaching, Atlanta, GA

National Association for Research in Science Teaching 2018
Conference Paper

Abstract
The transition of the science education community toward meeting the call of *The Framework for K-12 Science Education* (NRC, 2012) and *Next Generation Science Standards* (NGSS Lead States, 2013) requires science teachers to attend to the diverse backgrounds and abilities of their students. We describe our process for designing NGSS-aligned instructionally supportive assessment tasks for classrooms with diverse student populations in middle school Life and Physical Sciences. Some of the design features of tasks include familiar and relevant science phenomena, language supports, and various levels of scaffolding that support the integrated use of science and engineering practices, crosscutting concepts, and disciplinary core ideas. Our intentional effort to create assessment tasks that elicit the three dimensions of science proficiency is meant to provide equitable access and reduce barriers for a wide range of students, particularly linguistically and culturally diverse student populations.

Introduction

The Framework for K-12 Science Education (NRC, 2012) and *Next Generation Science Standards* (NGSS) (NGSS Lead States, 2013) emphasize a vision for learning whereby students engage in authentic science in ways that should be meaningful, relevant, and active. One significant challenge facing science teachers bringing the Next Generation Science Standards into their diverse classrooms involves developing new ways to accurately assess 3-dimensional learning - where such information is helpful for students in making meaningful progress towards a targeted scientific topic. In this three-dimensional vision of science learning, disciplinary core ideas, scientific and engineering practices, and crosscutting concepts are integrated as the means to make sense of natural phenomena and solve science and engineering problems. This is a purposeful and significant shift away from focusing on science content and science practices as disparate, one-dimensional learning goals. Accordingly, the new standards – called performance expectations (PEs) – are written as three-dimensional statements that describe the proficiencies associated to a particular scientific concept that students must achieve by the end of a grade band. For all students to be proficient in the NGSS performance expectations, they will need a broad range of opportunities in their science classrooms to provide space for demonstrating their developing proficiencies. Well-designed classroom-based assessments are critical as they often serve as a guide for teachers and students as they address targeted concepts during instruction. Unlike traditionally-designed assessments, NGSS-aligned assessments must not only provide information about student proficiency, they must also bridge patterns of inequity and appropriately assess a wide range of students. Importantly, assessments for NGSS must be accessible to students from a wide range of backgrounds and experiences (Pellegrino et al., 2014), which is reflected in our nation's increasingly diverse schools.

In this paper, our goal is to describe our approach in using Evidence Centered Design (ECD) to develop instructionally supportive assessment tasks for diverse students in classrooms that are implementing NGSS. ECD provides a structured and principled approach to assessment task design and the development of instructionally supportive assessment tasks. By *instructionally supportive*, we mean tasks that can be integrated into instruction to support teachers in learning how students are progressing toward the targeted PEs. These tasks are intended to be used formatively; meaning, they are meant to inform the direction of instruction, ensuring attention to how diverse students relate and respond to assessment tasks. Beyond this, we focus our design work on meeting the needs of both teachers and students in school settings that serve students from diverse populations. In our work, we strive to create equitable instructionally supportive assessments by recognizing, identifying, and addressing systemic bias. Today's classrooms include students that are diverse in a myriad of ways. These students also bring a vast range of experiences, background knowledge, and abilities to the classroom. Hence, it is imperative that science assessments be made accessible to all students. One challenge in developing instructionally supportive assessment opportunities for diverse student populations that aid teachers in effectively collecting, synthesizing, and interpreting information regarding their students' developing proficiencies is designing tasks that will not introduce additional challenges for these students.

Our design work takes place in a multi-institutional collaborative, called the Next Generation Science Assessment (NGSA) group (Harris et al., 2016). The NGSA group has recognized a need for a clear and efficient approach to designing assessment tasks so that teachers can support their students in making progress toward achieving the PEs. An important part of the design process is ensuring that the tasks are designed with equity and fairness in mind.

We begin our paper with a background on science assessment and equity, and then describe how we incorporate equity and fairness throughout our design approach. We provide example cases of tasks to illustrate how we apply an equity and fairness framework. We conclude by briefly considering implications of our work for creating NGSS-aligned assessment tasks for use in diverse classroom settings.

Background: Science Assessment and Equity

According to NGSS, being proficient in science requires students to integrate all three dimensions as intertwined knowledge that builds over time, which differs from traditional science content acquisition requirements (NRC, 2012). The way students build their science knowledge can depend on the kinds of perspectives students bring with them into the classroom (NRC, 2012). The diversity in perspectives often comes from cultural heritage, religious beliefs, and family background, which may create challenges for students whose experiences may be different from the dominant, and more widely accepted, "ways of knowing" (Brown & Abell, 2007). As students work to build their proficiency towards NGSS PEs, they must navigate their existing understandings and experiences of the natural world as they relate to the scientific phenomena they encounter in NGSS. In addition, they must successfully demonstrate this navigation in their assessments. Teachers are also faced with planning instruction that attends to how students' progress towards PEs, while taking into account background and perspectives (NRC, 2015).

The National Research Council's report on NGSS assessment (Pellegrino et al., 2014) notes that effective assessments must allow for the diverse ways in which students may express their developing understanding of science. In order for all students to achieve the NGSS performance expectations, all students need equitable and fair opportunities to learn and demonstrate their working knowledge of science. Since today's classrooms include students that are diverse ethnically, linguistically, socioeconomically, and in learning and physical ability, and since students also bring a range of experiences, background knowledge, and abilities to the classroom (Brown & Ryoo, 2008; Calabrese Barton, 2001; Lee, 2001), science assessments should be accessible to all students by attending to a variety of student needs. We describe accessible assessments as assessments that reduce bias which may unfairly disadvantage certain students over other students. When assessments are designed from a "one size fits all" perspective (Borthwick-Duffy, Palmer, & Lane, 1996), many students' needs are not acknowledged nor addressed. Assessment design that takes a principled and systematic approach can provide the necessary structure that assures that each task not only addresses all three dimensions of NGSS, but also attends to the diverse experiences and needs of students at each step of the design process.

Figure 1 shows the present diversity in schools across the nation and the ways task design can address the needs of various student demographics. The top level of the figure illustrates our nation's school diversity that assessment designers must attend to as they engage in task design. Drawn from NGSS Appendix D, we identify six non-dominant demographic groups - economically disadvantaged students, ethnically diverse students, students with disabilities, English language learners (ELL), gifted students, and girls and young women. The second level shows two critically important supports for meeting the needs of diverse learners – student engagement and language supports. Both have been identified as important for providing effective learning opportunities for the identified groups (see, for example, CAST, 2013; Luykx, Lee, Mahotiere, Lester, Hart, & Deaktor, 2007; Calabrese Barton, 2001). For example, anchoring assessment tasks in relevant scientific phenomena and practices can increase engagement in science learning activities (Calabrese Barton, 2001; Januszyk, Miller, & Lee, 2016). In addition, providing language supports to more broadly address how students communicate through reading and writing (e.g., English language learners, students with disabilities, and students from socioeconomically disadvantaged populations) can increase accessibility to students during science instruction (Siegal, 2007).

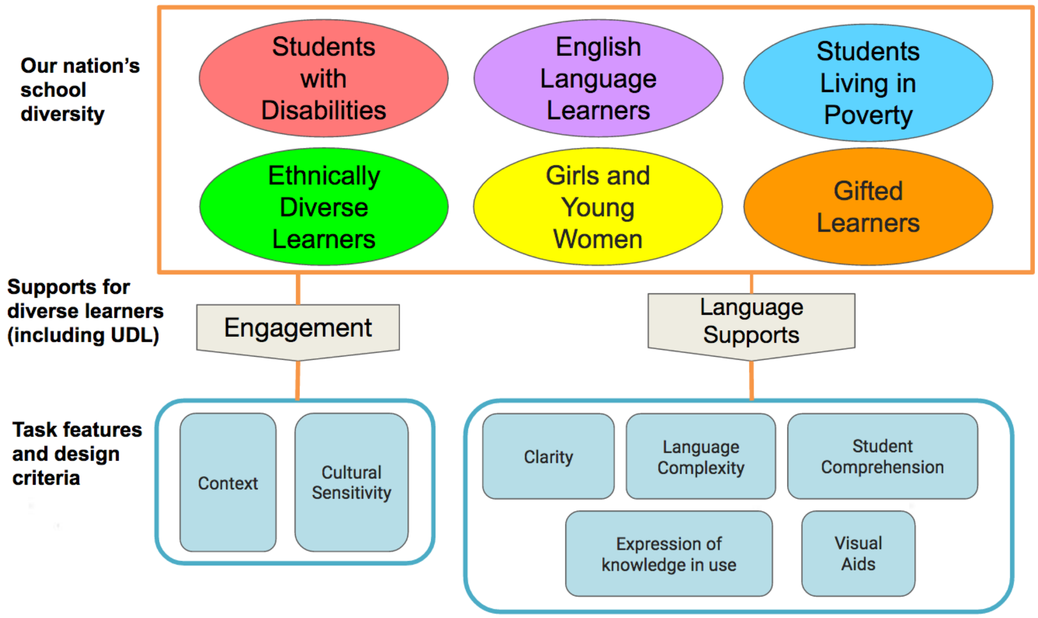


Figure 1: Representation of Diversity in Schools and Associated Task Features

Engagement. Universal Design for Learning (UDL; (Rose & Meyer, 2006) provides guidance for the development of flexible learning environments that can accommodate individual learning differences. One important UDL principle emphasizes "Multiple Means of Engagement" (CAST, 2011). This principle emphasizes the "why" of learning; essentially, the aspects of learning that motivate students to continue to learn. Students are engaged by information and activities that are relevant and familiar to their interests. However, we do not assume that all students will find the same activities or information equally relevant, familiar, or valuable. To approach an equitable learning opportunity for students, it is critical to provide options that optimize what is relevant, valuable, and meaningful to the learner while aligning to NGSS performance expectations. UDL categorizes engagement to include cultural sensitivity and context. In our design process, cultural sensitivity takes into account elements of the task that refer to gender, race, ethnicity, religion, geographical region. For example, Buck, et al. (2002) found advanced placement tests showed that students did better on questions related to content areas that highlight or identify their own gender.

We also consider how the context of a question can affect how a student interprets, relates to, and responds to a task question (McCullough, 2004). According to Solano-Flores and Nelson-Barber (2000), students' background and experiences outside of the formal school learning have a significant influence on how they interpret and respond to science assessment items. Since all learners need to be able to generalize and transfer their learning to new contexts

(CAST, 2011), our task development process utilizes a variety of scenarios with the goal of placing the disciplinary core ideas and practices in familiar settings and activities for students.

Language. Crafting tasks with special attention to language can create or remove barriers for students. Januszyk, Miller, and Lee (2016) argue that engagement in science and engineering practices requires language and discourse support, especially for ELL students and students with limited literacy development. One such support is the bridging of the cultural divide between learning to use the language of science and maintaining cultural identity (Brown and Ryoo, 2008). ELLs and students unfamiliar with science content interpret science terms with reference to their everyday meanings rather than scientific meanings (Luykx et. al, 2007). Our task development process identifies ways to support comprehension of the task. Providing multiple representations such as illustrations, simulations, images or interactive graphics can make the information in text more comprehensible for any student and accessible for some who would find it inaccessible in text (CAST, 2011). These supports can also help students bridge science language and their cultural identities.

The third level of Figure 1 presents important task design features and criteria that designers should attend to in their development work. Issues related to equity and diversity become even more important when standards are translated into instructional materials and assessments (NRC, 2014). Without thoughtful attention to the diverse needs of all students, assessments may unintentionally create unfair obstacles and limit accessibility for such groups of students as English language learners, students from racial and ethnic groups, and students with disabilities. Therefore, it is critical to develop tasks that increase access to the constructs being measured and that allow greater opportunities for students to demonstrate their proficiency in science so that teachers could effectively evaluate, and provide the necessary learning supports for students. We created a task design and development structure that 1) integrated culturally relevant phenomena (Calabrese Barton, 1998, 2001; Brown & Ryoo, 2008; Mueller & Bentley, 2009) into the extensive analysis of the DCIs as a way to bridge students' home and school lives, 2) used intersecting science practices and crosscutting concepts to provide various opportunities for students to demonstrate proficiency at intermediate points of instruction, 3) used research on cultural biases, and linguistic and cognitive demands in assessments (Luykx, et al., 2007; Solano-Flores & Nelson Barber, 2000; CAST 2011) to identify specific task design and development features for the creation of task design specifications, and 4) systematically reviewed each task as they were transferred to an online portal. This goal of this structure is to provide a principled way to integrate task features that move towards providing equitable opportunities for students to demonstrate their developing proficiency in NGSS PEs.

How We Design for Equity and Fairness

Our NGSA approach draws from evidence-centered design (ECD) (Mislevy & Haertel, 2006), which has gained widespread attention as a comprehensive approach for principled assessment design. ECD emphasizes the evidentiary base for specifying coherent, logical relationships among the (a) learning goals that comprise the constructs to be measured (i.e., the claims articulating what students know and can do); (b) evidence in the form of observations, behaviors, or performances that should reveal the target constructs; and (c) features of tasks or situations that should elicit those behaviors or performances. In terms of designing for the NGSS, ECD provides a structure for analyzing the three dimensions of NGSS and specifying their essential and assessable components. It also provides systematic and consistent opportunities for the integration of task features that support students from diverse backgrounds; making assessment tasks more accessible. By *accessible*, we mean that tasks provide students with opportunities to demonstrate science knowledge, skills, and abilities, through a variety of appropriate methods and procedures for different students (Suskie, 2000) and through the acknowledgment of sociocultural influences, such as values, beliefs, experiences, communication patterns, teaching and learning styles, socioeconomic conditions, and epistemologies inherent in students' backgrounds and cultures (Solano-Flores & Nelson-Barber, 2001; Lee, 2003).

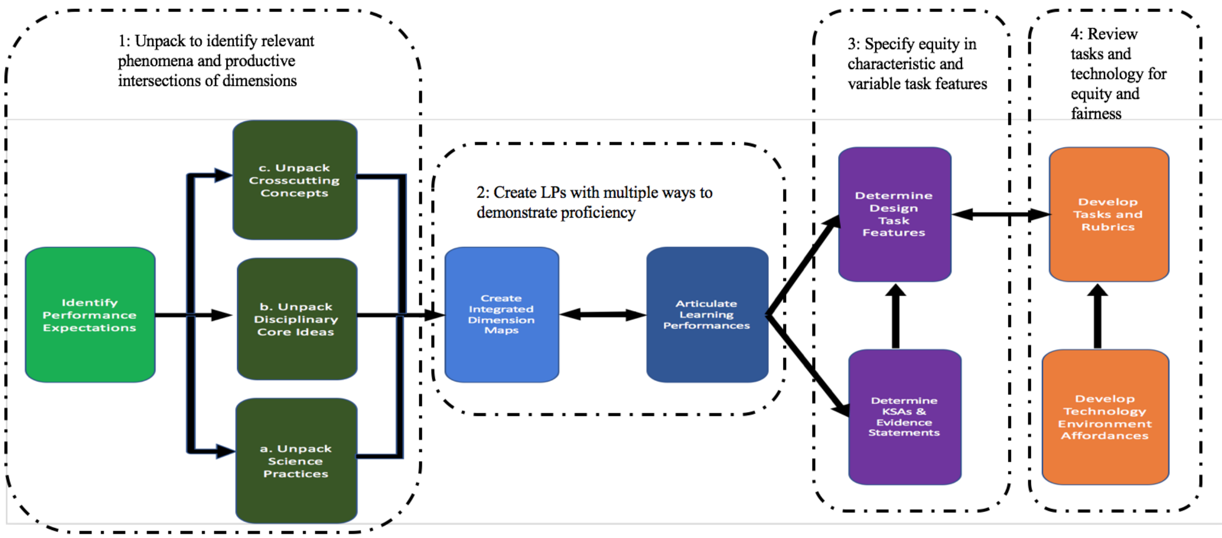


Figure 2: Integrating Equity Considerations in Our Task Design Approach

Throughout our NGSA task design process (Harris, Krajcik, Pellegrino, & McElhaney, 2016), we strive to give attention to providing equitable opportunities for students to demonstrate their progress. Figure 2 illustrates how we incorporate equity as we move through 4 key steps of the process. In our first step, typically referred to as *domain analysis* in ECD, we unpack the three NGSS dimensions in the target performance expectations. As part of the unpacking, we identify phenomena that are both relevant to the disciplinary core ideas and likely of relevance to students, specify student background knowledge and/or abilities for all three dimensions, and explicate productive intersections and integration of dimensions for demonstrating proficiency. In our second step, we construct learning performances. A learning performance (LP) is a three-dimensional statement that is smaller in scope and represents partial coverage of a performance expectation. Each learning performance describes an essential part of a performance expectation that students would need to achieve at some point during instruction to ensure that they are progressing toward achieving the more comprehensive performance expectation. A set of learning performances would provide complete coverage of one performance expectation. In our process, we create multiple learning performances for a performance expectation with different ways for students to demonstrate proficiency; enabling instruction to work with various students' strengths and abilities. In the third step, we specify equity in characteristic and variable task features. Characteristic task features describe the attributes that are common across all the tasks of a learning performance. Variable task features describe the features that can vary across tasks, such as the level of scaffolding that varies task difficulty. Finally, in our fourth step, we conduct a review of drafted and final versions of tasks that are designed in an online technology portal. Our review process uses a set of design criteria to determine whether each task meets certain design requirements, such as grade appropriate reading levels and a task scenario that is relevant and engaging.

Below, we describe the four steps in greater detail and in the context of our ongoing work in designing assessment tasks for middle school life science and physical science instruction (see Figure 2). In our physical science design work, we focus on the disciplinary core idea of *Matter and Its Interactions* because it helps answer two critical everyday questions about the physical world: "What is everything made of?" and "Why do things happen?" (NRC, 2012). In life science, we focus on two of the four substantive ideas in middle school life science: *Structures and Processes of Organisms* and *Interactions, Energy and Dynamics of Ecosystems*. We selected these two core ideas because they help answer two critical questions in biology: "How do organisms obtain and use the matter and energy they need to live and grow?" and "How do matter and energy move through an ecosystem?" Table 1 shows our selected performance expectations in physical science and life science for middle school.

Table 1: Selected Physical and Life Science Performance Expectations for our NGSA Work

Physical Science Performance Expectations	Life Science Performance Expectations
MS-PS1-2. Analyze and interpret data on the properties of substances before and after the substances interact to determine if a chemical reaction has occurred.	*MS-LS1-6.* Construct a scientific explanation based on evidence for the role of photosynthesis in the cycling of matter and flow of energy into and out of organisms.
MS-PS1-4. Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.	*MS-LS1-7.* Develop a model to describe how food is rearranged through chemical reactions forming new molecules that support growth and/or release energy as this matter moves through an organism.
MS-PS1-5. Develop and use a model to describe how the total number of atoms does not change in a chemical reaction and thus mass is conserved.	*MS-LS2-2.* Construct an explanation that predicts patterns of interactions among organisms across multiple ecosystems.
MS-PS3-4. Plan an investigation to determine the relationships among the energy transferred, the type of matter, the mass and the change in the average kinetic energy of the particles as measured by the temperature of the sample.	*MS-LS2-3.* Develop a model to describe the cycling of matter and flow of energy among living and nonliving parts of an Ecosystem.

1. *Unpack to identify relevant phenomena and productive intersections of dimensions*

To create assessment tasks that explicate key ideas of each dimension of the NGSS performance expectations (PEs), we begin the process with what we call *unpacking* (Harris et al., 2016). In unpacking, we conduct a careful and thorough domain analysis of each of the NGSS dimensions that are encompassed in the target PEs to ensure that assessment designers have a deep understanding of these statements. In the unpacking of the disciplinary core ideas, we elaborate the meaning of key sub-ideas, define clear expectations for what ideas students would be expected to use, demarcate boundaries for what students are or are not expected to know, identify background knowledge that is expected of students in order to develop a grade-level-appropriate understanding of aspects that correspond to a disciplinary core idea, and identify research-based problematic student ideas. We also identify phenomena that provide compelling examples of the disciplinary core idea. Identifying science phenomena that can help bridge student cultural and lived experiences with science is particularly significant because it may facilitate student understanding of traditionally challenging concepts (Calabrese Barton, 1998; 2001; Fries-Gaither, 2009). The process of unpacking DCI provides a suitable structure for incorporating relevant phenomena as it allows for multiple entry points for students. Specifying relevant phenomena can potentially help bridge students' everyday and school lives, while still meeting the expectations of the NGSS (NRC, 2012). Such phenomena utilize the backgrounds, knowledges, and experiences of the students to inform how tasks should be designed. Thus, as part of our unpacking of terms and concepts related to DCIs, we consider a range of real-world phenomena from diverse cultures and geographical regions that may serve to aid students in eliciting their understanding of these concepts as not merely abstract terms, but ideas that could be tangible and relevant to their own lives. Identifying examples in the unpacking phase then translates to our specification of task design features for each task (see Figure 2, step 3). Furthermore, we carefully identify and revise English language that is amenable to various linguistic backgrounds. For example, words that might change meaning with context are excluded when including phenomena that could help students identify with the disciplinary core idea.

When we unpack the science practices, we articulate specific aspects of practices students are to perform, specify the evidence required for students to demonstrate a high level of proficiency with a practice, identify prior knowledge that is required of students to demonstrate the practice, and identify common challenges that students may encounter as they are developing sophistication with the practice. We also identify productive intersections between the practice and other science practices. Identifying intersections allows us to ensure that there are multiple ways that students can demonstrate their proficiency with the performances encompassed in the PEs and LPs.

As part of unpacking the crosscutting concepts, we identify the important aspects of each, as well as how the crosscutting concepts intersect with targeted science practices and within a particular set of disciplinary core ideas. Similar to our unpacking of practices, we also specify the evidence required for a student to demonstrate a high level of proficiency with the crosscutting concept.

2. *Create LPs with multiple ways to demonstrate proficiency*

Our approach consists of articulating 3-dimensional claims called learning performances (LPs) that describe the intermediate performances that students need to achieve to progress toward a PE. Table 2 shows five learning performances we created from the unpacking of PE MS PS1-4 for physical science, and five created for PE MS-LS1-6 for life science. A set of LPs are intended to work together to illustrate how a student might achieve proficiency of the PE. They function together to inform teachers about students' progress over time. In our design process, the LPs provide varied ways for students to demonstrate achievement of different aspects of the performance expectation. In science instruction, inclusive instructional strategies include a range of techniques and approaches that build on students' interests, abilities, and backgrounds in order to engage them more meaningfully and support them in sustained learning (NRC, 2012). We applied this principle in our development of LPs. To avoid a "one size fits all" approach to assessment design, we diversified the science practices in the LPs underlying each PE. Varying the science practice provides increased opportunities for more students to demonstrate proficiency towards the PE.

This variation among LPs will also create diverse learning opportunities through instruction that attends to students that might show various interests or abilities. The formative nature of the tasks lend themselves to instruction that works with a student's abilities and helps build on challenging aspects of the student's growing proficiency.

Table 2: Learning Performances Aligned to MS PS1-4 and MS LS1-6.

Learning Performances for Physical Science MS PS1-4 *Develop a model that predicts and describes changes in particle motion, temperature, and state of a pure substance when thermal energy is added or removed.*
LP E-01: Evaluate a model that uses a particle view of matter to explain how states of matter are similar and/or different from each other.
LP E-02: Develop a model that explains how particle motion changes when thermal energy is transferred to or from a substance without changing state
LP E-03: Develop a model to explain the change in the state of a substance caused by transferring thermal energy to or from a sample.
LP E-04: Use evidence from a simulation to construct a scientific explanation about how the average kinetic energy and the temperature of a substance change when thermal energy is transferred from or to a sample.
LP E-05: Develop a model that includes a particle view of matter to predict how the average kinetic energy and the temperature of a substance change when thermal energy is transferred from or to a sample.
Learning Performances for Life Sciences *MS-LS1-6* *Construct a scientific explanation based on evidence for the role of photosynthesis in the cycling of matter and flow of energy into and out of organisms.*
LP P-01: Analyze and interpret data to determine whether plants and other photosynthetic organisms grow with the input of energy from sunlight.
LP P-02: Analyze and interpret data to determine whether plants and other photosynthetic organisms take in water, carbon dioxide, and (e.g., sunlight), to produce food (sugar) and oxygen.
LP P-03: Develop a model that shows that plants (or other photosynthetic organisms) take in water and carbon dioxide to form food (sugar) and oxygen.
LP-04: Evaluate how well a model shows that plants and other photosynthetic organisms use energy from the Sun to drive the production of food (sugar) and oxygen.

LP P-05: Construct a scientific explanation for how plants (and other photosynthetic organisms) are able to use energy from light and matter from the sugar they produce to grow and support their other necessary (life--supporting) functions (explanations should attend to matter and energy).

3. *Specify equity in characteristic and variable task features*

Leveraging the unpacking of science practices, crosscutting concepts, and disciplinary core ideas and the articulation of LPs, we then move towards the creation of task design specifications, called *design patterns*, for each LP. Design patterns are detailed outlines of what to assess in a particular LP and how to assess it. Here, we also specify the focal knowledge, skills, and abilities (FKSAs) that students need in order to demonstrate the learning performance, what constitutes as evidence of a student having mastered the FKSA, and task features that would prompt students to demonstrate proficiency of the learning performance.

Task features are designed to help the student access FKSAs when working through an assessment task. However, if task features are inadequately designed, they can introduce additional complexity and/or bias to the task, thereby decreasing access to the FKSAs being measured (Abedi & Lord, 2001; Ahmed & Pollitt, 2007; Cooper & Dunne 2000; Solano-Flores & Nelson-Barber, 2001; Solano-Flores & Trumball, 2003). To address this concern, we identify characteristic and variable features that can support different groups of students. For example, to ensure the engagement of students from non-dominant groups (see Figure 1), a characteristic task feature might be that all tasks must relate to relevant, real-world context (Barton, 2001; Lee, 2001) that students can relate to or are familiar with. Similarly, a variable task feature might be that a task provides a high level of scaffolding for the student to answer the FKSA, versus a task that provides little or no scaffolding. The level of scaffolding gives teachers instructional options as they gather information about students' progress towards the PEs.

Table 3 shows a general set of characteristic and variable features that create various assessment opportunities for diverse student populations. The features are categorized by student engagement and language practices as identified by our research. Note that our design patterns (not shown here) contain more detailed guidelines for each feature.

Table 3: Characteristic and Variable Task Features to Promote Equitable Assessment Opportunities.

Characteristic Task Features
Student Engagement **Item Engagement and Contextualization:** Task scenarios contain familiar, relevant, or authentic situations *Vivid Phenomenon-based Situations* • Tasks elicit prior knowledge, the use of relevant phenomena, and representations (Stern & Roseman, 2004) • Tasks use modeling and representations (Lee et al, 2013) • Tasks explicitly focus on lived experiences (Barton et al., 2008) • Tasks connect school culture, students' home communities, and real-world contexts (Barton, 2001; Lee, 2001; Ostergaard et al., 2010) *Relating Students and Nature*

- Tasks provide ecological awareness and encourage relationships with and awareness of nature (Mueller & Bentley, 2009).

Cultural Sensitivity: Task scenarios reduce bias and stereotypes for a particular gender, race, SES, or geography

- Tasks are sensitive to how of cultural, gender, SES, race, and religion are mentioned or discussed
- Tasks include a task scenario that is inclusive (i.e., not representative of a particular group, such as higher SES)

Language Supports

Clarity of Prompt: Task prompts must clearly elicit the target performance and avoid superfluous information

Language Complexity: Tasks must be accessible to students with limited English reading ability

- Tasks use everyday language, particularly for ELL (Ryoo, 2015)
- Tasks use vocabulary words, sentence structure, and reading level that are grade level appropriate
- Tasks use visual representations to help support vocabulary

Student Comprehension: Information in tasks is consistent and relevant throughout the task

- Tasks support prior knowledge, skills, and/or abilities
- Information in the tasks are organized in a manner that supports reading comprehension (e.g., labeled tables or diagrams and pictures, chunked information)
- Terms are used consistently throughout the task stem and prompt

Variable Task Features

Student Engagement

Student Response/ Expression of "Knowledge-in-Use": Tasks provide for varied modes of response

- Tasks use different response templates and/or scaffolds to support student expression of knowledge-in-use

Language Supports

Visual Aids: Tasks differ in the type of visual aids to support comprehension

- Tasks use interactive simulations as appropriate to aid comprehension (Quelmaz et al., 2012; Grotzer et al., 2014)

4. Design Cases: Review tasks and technology for equity and fairness

We iteratively refine each task using an elaborated version of the design framework. In step 4, we perform a comprehensive review where we evaluate the structure and content of the tasks to ensure increased likelihood of accessibility for diverse students. In our evaluation, we used a set of criteria, based on the characteristic and variable task features, with indicators, to determine whether the tasks sufficiently incorporate equitable task features. In our process, a team of reviewers completed 1 day of training on how to apply the criteria to the tasks, and 2 days of opportunities (intermittent) to discuss the review process and the tasks. Tasks were randomly assigned to reviewers and reviewers provided detailed feedback on each task. Finally, tasks were refined to address the feedback given during the review process.

Below, we present a design case to show how we used ECD to incorporate opportunities for equitable task design and development. In this design case, we show the important equity and fairness features of a life science task addressing the interactions of organisms in an ecosystem.

Design Case

Life Science Task with Explicit Attention to Features That Promote Equity

Barbara learned in school that coral reefs have animals called corals and small plant-like organisms called algae that live there. Algae live inside corals and make them look colorful. She also learned that rising ocean temperatures are harming **some** corals in coral reefs.

Barbara visited coral reefs the following summer. When she dove under the water, she saw a pattern where some corals looked **colorful** and **healthy** while others were **white** and **dead**.

Colorful and healthy coral reef

License: Public Domain

White and dead coral reef

Credit: Acropora
License: CC-BY 3.0

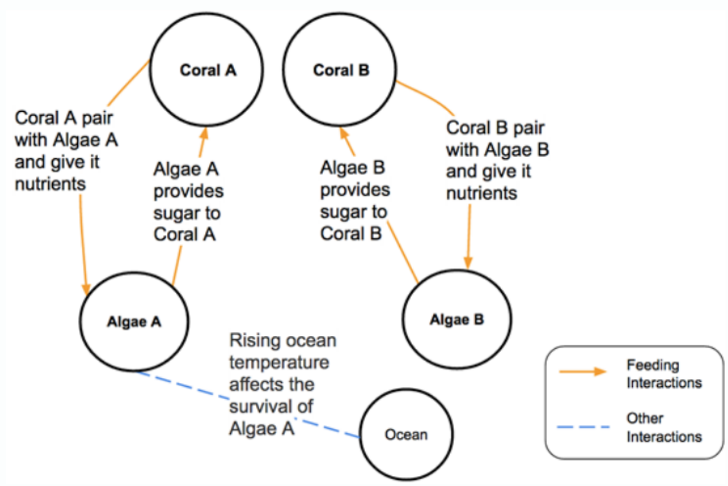
Barbara created a model using what she knew about interactions between ocean temperatures, corals, and algae. She wants to use her model to try and understand the pattern she observed when she visited coral reefs in the summer.

Question #1

Use Barbara's model to describe why some corals appear colorful and are healthy while other corals appear white and dead.

Be sure to include what you know about mutually beneficial relationships and the role of algae and rising ocean temperatures in your response.

Type answer here



This life science task incorporates a range of characteristic and variable features that focus on student engagement and language supports.

Characteristic Task Features for Student Engagement

Item engagement and context: In the unpacking of the DCI, science phenomena that can help bridge students' everyday experiences with relevant science phenomena were identified and applied to this task. Relevant phenomena can potentially help bridge students' everyday and school lives, while still meeting the expectations of the NGSS. In this task, the character in the task, Barbara, is a student learning about coral reefs at school. This is similar to students that may be learning about organisms that reside in various ecosystems. The task extends the experience of the student by creating a story where Barbara visited coral reefs during her summer break. This extension can help the student connect to nature in ways that they may not be available in other tasks. The task includes real-world phenomena from distal geographical regions which may aid students in eliciting their understanding of these concepts as ideas are tangible and relevant to their own lives.

Cultural sensitivity: The task scenario of underwater organisms may not be familiar to all students. Therefore, the task contains pictures and definitions. These additional supports help reduce bias towards students who might live in urban or rural communities, or in regions that are distant from large bodies of water, while still providing real-world science phenomena that directly align to the NGSS DCI.

Characteristic Task Features for Language Supports

Clarity of prompt: In this task, additional information about mutually beneficial interactions is provided to support the students as they respond to the question. An additional revision to consider for this task may be to bullet the additional information to guide the student's focus more closely.

Language complexity: The task uses everyday language to help bridge familiar language for students and scientific language. To help students who may still have difficulty utilizing science discourse as they work with science phenomena, the task uses simple and consistent language with concise sentences. The task also includes an active voice to support student engagement.

Student comprehension: As students work through the task, the task provides connections between the text and the images in order to help the student make sense of the storyline. Supports that organize information in the task, such as diagrams, pictures, and chunked information, help support executive functions skills for students that may struggle with managing, organizing, and remembering information in the task (CAST, 2015). As mentioned in cultural sensitivity, all students may not be familiar with coral reefs. Because of this, the task provides additional information that explains what coral reefs are and what grows in them. These descriptions are connected to the images of different coral reefs (healthy and dead), as described in the task scenario, and will help the students work with the information provided in the model.

Variable Task Features for Language Supports

Visual Aids: Students that vary in cognitive ability, have limited literacy skills, and ELL students, can benefit from visual aids that support understanding of science phenomena. The task provides students with different images of coral reefs and their locations to help the student make sense of the science phenomena.

Conclusions and Implications

The design and development of three dimensional tasks is challenging but necessary, as most current assessments cannot be used to measure the increased range of activities and interactions that are required by NGSS. With increased diversity in classrooms, teachers must work with students that may not respond well to traditionally-designed assessment tasks. The use of evidence-centered design as a foundation for our work provides us with a principled approach for embedding research-based ideas into the assessment design process while considering diverse students. An important decision for us has been the concerted attention to equity throughout the design process. Our design process is designed to minimize construct-irrelevant variance and reduce barriers so that students from diverse backgrounds can fully demonstrate their proficiency in NGSS PEs. Our task design process offers a useful structure that appears to help designers build coherence for equity and diversity in the construction of tasks. This design process works to address and attend to cultural bias and their effects on students from non-dominant groups. We use engagement and language as the main focus when working to design and develop tasks through ECD. Attending to student relevance through unpacking, formulating LPs that provide multiple ways of demonstrating proficiency, and creating design patterns that address student engagement and language use, serves as the foundation for a coherent task design process promoting equity.

For practitioners, we believe the benefit lies in providing teachers with tasks that are more accessible for students and useful for garnering information about how a broad range of students are developing proficiencies with the targeted NGSS PEs in both physical science and life sciences. Education practitioners may efficiently access our tasks to integrate in their existing science curriculum materials. Students from diverse backgrounds may be able to demonstrate their developing proficiency towards performance expectations with less bias than traditional formative and summative assessments. This may allow for instruction that is more relevant to students' lives. Our work provides opportunities for a more informed teacher that can mitigate implicit assumptions about non-dominant students' developing proficiencies and utilize students' diverse knowledge about the world during instruction and learning that progresses in relevant and meaningful ways to students.

We contribute to the body of work in equity and diversity and instruction by providing an innovative approach for developing assessment tasks to be used formatively in diverse classrooms. This approach has 4 key steps: (1) Unpack to identify relevant phenomena, student background knowledge and/or abilities, and productive intersections of dimensions, (2) Create LPs with multiple ways to demonstrate proficiency, (3) Specify equity in characteristic and variable task features, and (4) Review tasks and technology for equity and fairness. We have found that the careful specifying of characteristic and variable task features – prompt clarity and comprehensibility, item engagement and contextualization, language complexity, cultural sensitivity, student response and expression of knowledge, and visual aids – help minimize construct-irrelevant variance, such as including unfamiliar vocabulary and using highly inferential language, that can impede on all students' ability to successfully complete tasks. At the same time, the task features provide supports for ELL students and students who may have limited reading levels, in addition to dominant student groups.

Although our work is ongoing, outcomes from our work include design patterns that inform task development, an exemplar set of formative assessment tasks, and a set of task review criteria that are used in the iterative evaluation of these tasks. The task review criteria are used in a review process that focuses on (a) using clear, simple, grade-appropriate language, with little ambiguity, (b) decreasing the cognitive load for students that may struggle with executive functioning, (c) providing clarity in the task prompt, (d) contextualizing the disciplinary core ideas in familiar scenarios for a wide range of students, and (e) identifying and correcting inaccuracies in the task, with the overall aim to make all tasks usable in a range of classrooms. We will continue to refine and validate our design process to be more robust, systematic, and streamlined. We will also work towards validating our tasks with an eye towards equity. Furthermore, we are in the process of developing a guide for researchers and practitioners that will support them in the use of our tasks and rubrics, as well as continue to conduct classroom-based research to better understand how teachers can use these tasks effectively in diverse school settings.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219-234.

Ahmed, A., & Pollitt, A. (2007). Improving the quality of contextualized questions: An experimental investigation of focus. *Assessment in Education*, *14*(2), 201-232.

Airasian P.W.Empiricism and values: Two faces of educational change. *Studies in Educational Evaluation*, 27 (5) (1997), pp. 433-445

Barton, A. C. (1998). Teaching science with homeless children: Pedagogy, representation, and identity. *Journal of research in science teaching*, *35*(4), 379-394.

Barton, A. C. (2001). Science education in urban settings: Seeking new ways of praxis through critical ethnography. *Journal of Research in Science Teaching*, *38*(8), 899-917.

Borthwick-Duffy, S. A., Palmer, D. S., & Lane, K. L. (1996). One size doesn't fit all: Full inclusion and individual differences. *Journal of Behavioral Education*, *6*(3), 311-329.

Brown, P. L., & Abell, S. K. (2007). Project-based science. *Science and Children*, *45*(4), 60.

Brown, B. A., & Ryoo, K. (2008). Teaching science as a language: A "content-first" approach to science teaching. *Journal of Research in Science Teaching*, *45*(5), 529-553.

Buck, G., Kostin, I., & Morgan, R. (2002). Examining the relationship of content to gender-based performance differences in advanced placement exams. *ETS Research Report Series*, *2002*(2).

CAST (2011). Universal design for learning guidelines version 2.0. Wakeeld, MA: Author.

CAST (2015). Universal Design for Learning. Available at www.cast.org

Cooper, B., & Dunne, M. (2000). Constructing the 'legitimate' goal of a 'realistic' maths item. *Assessment: Social practice and social product*, 87.

Fries-Gaither, J. (2009). Citizen Science Projects: Everyone's a Scientist. *Middle Ground*, *13*(2), 28.

Gorin, J.S., & Mislevy, R.J. (2013). *Inherent measurement challenges in the next generation science standards for both formative and summative assessment.* Commissioned paper presented at the K-12 Center at ETS Invitational Research Symposium Invitational Research Symposium on Science Assessment, Washington, DC

Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & McElhaney, K. W. (2016). *Constructing assessment tasks that blend disciplinary core ideas, crosscutting concepts, and science practices for classroom formative applications.* Menlo Park, CA: SRI International.

Januszyk, R., Miller, E. C., & Lee, O. (2016). Addressing Student Diversity and Equity. *Science and Children*, *53*(8), 28.

Lee, O. (2001). Culture and language in science education: What do we know and what do we need to know? *Journal of Research in Science Teaching*, *38*(5), 499-501.

Lee, O. (2003). Equity for linguistically and culturally diverse students in science education. *Teachers College Record*, *105*(3), 465-489.
Lee, O., & Luykx, A. (2007). Science education and student diversity: Race/ethnicity, language, culture, and socioeconomic status. *Handbook of research on science education*, 171-197.
Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J., & Deaktor, R. (2007). Cultural and home language influences on children's responses to science assessments. *Teachers College Record*, 109(4), 897-926.
McCullough, L. (2013). Gender, context, and physics assessment. *Journal of International Women's Studies*, *5*(4), 20-30.
Mislevy, R., & Haertel, G. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6–20.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
National Research Council (2014). *Developing Assessments for the Next Generation Science Standards*. Washington, DC: National Academies Press.
National Research Council (2015).*Guide to Implementing the Next Generation Science Standards.* Washington, DC: National Academies Press.
NGSS Lead States (2013). *Next Generation Science Standards: For states, by states.* Washington, DC: The National Academies Press.
Noble, T., Suarez, C., Rosebery, A., O'Connor, M. C., Warren, B., & Hudicourt-Barnes, J. (2012). "I never thought of it as freezing": How students answer questions on large-scale science tests and what they know about science. *Journal of Research in Science Teaching*, 49(6), 778-80
Noble, T., Suarez, C., Rosebery, A., O'Connor, M. C., Warren, B., & Hudicourt-Barnes, J. (2012). "I never thought of it as freezing": How students answer questions on large-scale science tests and what they know about science. *Journal of Research in Science Teaching*, 49(6), 778-803.
Pellegrino, J. W., Wilson, M. R., Koenig, J. A., & Beatty, A. S. (2014). *Developing Assessments for the Next Generation Science Standards*. National Academies Press. 500 Fifth Street NW, Washington, DC 20001.
Rose, D. H., & Meyer, A. (2006). A practical reader in universal design for learning. *Harvard Education Press*.
Solano-Flores, G., & Nelson-Barber, S. (2000). Cultural validity of assessments and assessment development procedures. In *annual meeting of the American Educational Research Association. New Orleans, LA, April*.
Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher*, *32*(2), 3-13.
Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching*, *38*(5), 553-573.
Siegel, M. A. (2007). Striving for equitable classroom assessments for linguistic minorities: Strategies for and effects of revising life science items. *Journal of Research in Science Teaching*, *44*(6), 864-881.
Suskie, L. (2000). Fair Assessment Practices Giving students equitable opportunities to demonstrate learning. *AAHE BULLETIN*, *52*(9), 7-9.

Thompson, S.J., Johnstone, C.J., and Thurlow, M.L. (2002). *Universal design applied to large scale assessments.* Synthesis Report 44. Minneapolis: University of Minnesota.
Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Acknowledgements
This material is based upon work supported by the Gordon and Betty Moore Foundation (Grant 4482) and by the National Science Foundation (NSF) (Grants 1316903, 1316908, and 1316874). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Gordon and Betty Moore Foundation and NSF.