Guidelines for the Evaluation of Instructional Materials in Science
Table of Contents

1. Introduction ............................................................................................................................................ 3
   Origin of the Guidelines .......................................................................................................................... 3
   Rationale for the Guidelines .................................................................................................................. 4
   About the Development of the Guidelines ............................................................................................... 5
   Project Participants ................................................................................................................................... 7

2. Criteria to Evaluate the Quality of Instructional Materials in Science ............................................ 9
   Category A: Instructional Materials Support NGSS-Driven Learning Goals ........................................ 11
   Category B: Instructional Materials Provide Coherence across the Three Dimensions ..................... 25
   Category C: Instructional Materials Support Learning Experiences across the Three Dimensions .... 31
   Category D: Instructional Materials Provide Ways to Monitor Learning across the Three Dimensions .. 40
   Concluding Remarks ............................................................................................................................... 46

3. Evaluation Systems for Assessing the Quality of Instructional Materials in Science .................. 47
   Guidelines for Evaluation Systems .......................................................................................................... 48
   Guidelines for Evaluation Tools ............................................................................................................. 52
   Guidelines for Evaluation Processes ....................................................................................................... 58
   Concluding Remarks ............................................................................................................................... 61

4. References .............................................................................................................................................. 62

Quick Links
   Table 2.1 Criteria for the evaluation of instructional materials in science ............................................ 10
   Table 3.1 Guidelines for assessing the quality of instructional materials ............................................... 47
Introduction

Origin of the Guidelines

The United States is a world leader in science and technology innovation. In most STEM industries, the United States remains at the cutting edge of scientific discoveries driven by innovative thinkers and a creative workforce (Xue & Larson, 2015). Despite these achievements, the US system of science education has struggled to adequately prepare all students for careers in science and citizenship beyond the K-12 classroom (Duschl, Schweingruber, & Shouse, 2007; National Research Council, 2012). Against this backdrop, the National Academy of Sciences (NAS), the National Research Council (NRC), the National Science Foundation (NSF), and others have offered calls to action and have embarked on large-scale and ambitious initiatives to improve STEM education and monitor progress in these improvement efforts over the past 10 years (see figure 1.1). Thus, science education in the United States is in the midst of a significant transition aimed at improving science learning for all students.

Figure 1.1. Timeline of benchmark calls to action and reports.

Building from the release of Taking Science to School (Duschl, Schweingruber, & Shouse, 2007) and Rising Above the Gathering Storm (National Academy of Sciences et al., 2007), the NRC developed A Framework for K-12 Science Education (Framework; National Research Council, 2012). In 2013 the NGSS Lead States unveiled the Next Generation Science Standards (NGSS; NGSS Lead States, 2013) based on the NRC Framework. These latter two documents outlined a contemporary vision for science education in the United States:

*The overarching goal of our framework for K-12 science education is to ensure that by the end of 12th grade, all students have some appreciation of the beauty and wonder of science; possess sufficient knowledge of science and engineering to engage in public discussions on related issues; are careful consumers of scientific and technological information related to their everyday lives; are able to continue to learn about science outside school; and have the skills to enter careers of their choice, including (but not limited to) careers in science, engineering, and technology.* (NGSS Lead States, 2013, p. 1)
In parallel, the NRC published *Successful K-12 STEM Education* (NRC, 2011) in response to a request from Representative Frank Wolf (VA) to identify highly successful K-12 STEM schools and programs. The report identified three overarching goals for K-12 STEM education—expand the number of students who pursue advanced STEM degrees and careers, expand the STEM-capable workforce, and increase STEM literacy for all students—and made recommendations for reaching these goals.

Congress followed the *Successful K-12 STEM Education* report by asking the NSF to identify ways to track national progress toward the recommendations in the report. The resulting publication, *Monitoring Progress toward Successful K-12 STEM Education* (Monitoring Progress; NRC, 2013), described 14 indicators as large-scale measures of the health of STEM education in US schools. NSF released a Dear Colleague letter in response to the *Monitoring Progress* report; this letter solicited proposals that addressed one or more of the 14 indicators.

BSCS responded to the NSF Dear Colleague letter with a proposal that addresses Indicator 4, “Adoption of instructional materials in grades K-12 that embody the *Common Core State Standards for Mathematics* and *A Framework for K-12 Science Education*” (NRC, 2013, p. 19). In particular, Indicator 4 identifies two tiers of data collection related to STEM instructional materials. The first tier would identify the most widely adopted materials in US schools, and the second tier—with regard to science instructional materials—would identify a set of criteria for determining the degree to which instructional materials embody the standards and promote the vision of science education described in the Framework. BSCS proposed to construct a set of guidelines for developing a national measure of the extent to which science instructional materials exemplify the Framework and the NGSS. The purpose of this project was to

- solicit advice from leading science educators on the essential criteria and important characteristics of measures for evaluating science instructional materials;
- engage science education leaders in building consensus, where possible, regarding these criteria; and
- produce a set of Guidelines, based on this consensus, for the development of a measure to evaluate science instructional materials.

**Responding to the call to monitor progress, the Guidelines provide recommendations for the evaluation of widely used science instructional materials.**

**Rationale for the Guidelines**

Instructional materials play a central role in science classrooms (Banilower et al., 2013). They reside at the intersection of standards, pedagogy, and assessment, often providing engaging and academically challenging learning experiences for students (Arzi & White, 2008; BSCS, 2009; Schmidt, 2003). Instructional materials are a medium for making academic disciplines accessible to learners and are often students’ first introduction to those disciplines.

Research has shown that instructional materials have a major impact on science teaching. Studies have found that teachers follow the structure of their textbooks (Arzi & White, 2008; Reiser, 2013). Novice teachers, in particular, lean on instructional materials a great deal (Grossman & Thompson, 2004). This underscores the reality that instructional materials are a primary driver for how science is taught and learned by millions of students in US classrooms (Bybee, 2013; NRC, 1999; Taylor, Gardner, & Bybee, 2009). In short, high-quality instructional materials represent a critical link between the NGSS and the science instruction that students experience in the classroom. As such, these materials demand considerable thought and careful design (Bybee, 2013; Schmidt et al., 2001; Taylor, Gardner, & Bybee, 2009).
It follows, therefore, that we should assess the alignment of instructional materials with the vision and goals we have established for teaching and learning. However, evaluations of instructional materials that have been conducted in the past indicate that widely used materials often do not meet the high standards set by science education reforms (e.g., Good, 1993; Kesidou & Roseman, 2002; Schmidt, McKnight, & Raizen, 1997; Stern & Roseman, 2004). These evaluations found that, while the comprehensive programs they evaluated did some things well, they frequently included far more content than can be covered in a single school year and tended to lack coherence necessary for conceptual learning by students (Duschl, Schweingruber, & Shouse, 2007; Krajcik, McNeill, & Reiser, 2008).

Education leaders have called for systems to monitor progress toward achieving this new vision in science education (NRC, 2013). There are countless instructional materials beyond textbooks and each likely meets some science reform standards well, falling short in other areas. Given the diversity of available instructional materials, a means of evaluating them is essential not only to inform our understanding in light of the NGSS but also to push the field toward designing quality materials that remedy past weaknesses.

Based on advice from national science education leaders, these Guidelines put forth criteria for evaluating widely used science instructional materials.

About the Development of the Guidelines

We interpreted the ultimate task of analyzing instructional materials as consisting of two components: (1) assessing the degree of alignment of instructional materials with the vision and goals of the NRC Framework and the NGSS and (2) assessing the likelihood of the materials supporting teachers and students in bringing the vision and goals to life. These two components can be loosely characterized as “alignment” and “quality”. In approaching this task, the project participants built on considerable work done over the last two decades on defining and measuring the quality of instructional materials (e.g., BSCS, 2007; Davis & Krajcik, 2005; Duschl, Schweingruber, & Shouse, 2007; Edelson, 2001; Kesidou & Roseman, 2002; Roseman, Stern, & Koppal, 2010). However, the vision of science teaching and learning presented by the NRC Framework and the NGSS has new elements that have not been part of previous assessments of instructional materials, for example, the explicit inclusion of science and engineering practices (SEPs), crosscutting concepts (CCCs), and disciplinary core ideas (DCIs). Therefore, an important part of our work in this project has been to identify and explain these new components to support assessment of them in instructional materials.

The Guidelines provide answers to the following questions:

1. What are the evaluative criteria for instructional materials that support the vision of science education described in the NRC Framework and the NGSS? and
2. How do we measure the extent to which these criteria are present in materials?

The evaluative criteria referred to in the first question offer an operational definition of the vision put forth by the Framework and NGSS specified for instructional materials. To address this question, we reviewed research on the characteristics of high quality instructional materials and convened a group of science education leaders to assist us in identifying the most critical criteria to include in the Guidelines. These criteria discussed in chapter 2.

Answering the second question requires specifying the necessary characteristics of a measure to evaluate
the extent to which the proposed criteria are present in instructional materials. To do this, we reviewed existing evaluative measures and identified nine key characteristics, which are discussed in chapter 3.

The Guidelines report was developed over two years involving 10 BSCS researchers and 18 leaders in science education from across the country. The team brought to the project expertise in the NRC Framework and the NGSS, tools and processes used to evaluate the quality of science instructional materials, curriculum development, and policy-making.

The multi-stage development process included early efforts focused on relevant research and literature, a Summit of science education leaders to generate initial ideas about the Guidelines, the development of the Guidelines, and two rounds of review, feedback, and revision of the Guidelines. In preparation for the Summit, BSCS staff worked with a group of advisors to complete summaries and a crosswalk of widely used tools and processes used to assess the quality of instructional materials, synthesize current research and literature describing the characteristics of high-quality instructional materials, and identify other leaders in science education as Summit participants. During the Summit, these leaders identified characteristics of measures to assess specified evaluative criteria for science instructional materials. BSCS used ideas generated during the Summit to develop an initial draft of the Guidelines report. The draft was reviewed by Summit participants. The revised report was reviewed by others who responded to a public call and targeted invitations to the larger science education community at conferences hosted by the National Association for Research in Science Teaching (NARST), National Science Teachers Association (NSTA), Council of State Science Supervisors (CSSS), and the American Educational Research Association (AERA). The process enabled us to build on earlier work in evaluation of instructional materials in science and to incorporate new ideas.

The development of the Guidelines was motivated by the goal of the NRC Monitoring Progress report—to conduct a review of widely used instructional materials at a national level. To conduct such a review will ultimately require the development and testing of a complete, fully specified process for evaluating instructional materials—the development of which was beyond the scope of this project. This project was designed to take an important, substantive step by identifying attributes of instructional materials that should be assessed in such a process as well as criteria for evaluating them. These Guidelines are designed to support others in creating tools and processes that can be used to conduct valid and reliable assessments of instructional materials.

Having developed these Guidelines to support the goals of the Monitoring Progress report, we are also mindful of the fact that people in many different roles conduct reviews and evaluations of instructional materials for many purposes. We expect that people will find these guidelines useful for other purposes, such as textbook adoption decisions, teacher professional development, and revisions to materials to enhance their alignment to the NGSS and overall quality.
**Project Participants**

Project participants include BSCS staff, advisors who served on the leadership team, Summit participants and reviewers, and those who provided feedback during the public comment phase of the project.

**BSCS Project Team**

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>April Gardner</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>Jody Bintz</td>
<td>Co-principal Investigator</td>
</tr>
<tr>
<td>Audrey Mohan</td>
<td>Co-principal Investigator</td>
</tr>
<tr>
<td>Lindsey Mohan</td>
<td>Consulting Research Scientist</td>
</tr>
<tr>
<td>Daniel Edelson</td>
<td>Executive Director</td>
</tr>
<tr>
<td>Lisa Carey</td>
<td>Project Coordinator</td>
</tr>
<tr>
<td>Lindsey Mohan</td>
<td>Consulting Research Scientist</td>
</tr>
<tr>
<td>Daniel Edelson</td>
<td>Executive Director</td>
</tr>
<tr>
<td>Lisa Carey</td>
<td>Project Coordinator</td>
</tr>
</tbody>
</table>

**Leadership Team**

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>George DeBoer</td>
<td>American Association for the Advancement of Science</td>
</tr>
<tr>
<td>Kathy DiRanna</td>
<td>K-12 Alliance at WestEd</td>
</tr>
<tr>
<td>Joseph Krajcik</td>
<td>Michigan State University</td>
</tr>
<tr>
<td>James Short</td>
<td>Gottesman Center for Science Teaching and Learning, American Museum of Natural History, Currently, Carnegie Corporation of New York</td>
</tr>
</tbody>
</table>

**Summit Facilitators**

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne Kennedy</td>
<td>Kennedy Education Collaboratory</td>
</tr>
<tr>
<td>Dennis Schatz</td>
<td>Pacific Science Center</td>
</tr>
</tbody>
</table>
Summit Participants

Aneesha Badrinarayan  
Achieve

Rodger Bybee  
Former BSCS Executive Director

Barbara Nagle  
Lawrence Hall of Science

Richard Duschl  
Penn State University

David Evans  
National Science Teachers Association

Thomas Keller  
Maine Mathematics & Science Alliance

Matthew Krehbiel  
Council of State Science Supervisors  
Currently, Achieve, Inc.

Michael Lach  
University of Chicago

Vicky Lamoreaux  
North Thurston Public Schools

Tatiana Lim-Breitbart  
Aspire California College Preparatory Academy

Ann Rivet  
Teachers College, Columbia University  
Currently, National Science Foundation

Heidi Schweingruber  
National Board on Science Education

Sean Smith  
Horizon Research, Inc.

Suzanne Wilson †  
University of Connecticut

†unable to attend Summit, but served as a Guidelines reviewer

BSCS Summit Participants

Brooke Bourdélat-Parks  
Senior Science Educator

Sue Kowalski  
Senior Research Scientist

Molly Stuhlsatz  
Research Scientist

Anne Westbrook  
Science Educator

Chris Wilson  
Senior Research Scientist

Project Evaluator

Dina Drits-Esser  
University of Utah
Criteria to Evaluate the Quality of Instructional Materials in Science

The NRC *Framework* and the NGSS now provide a foundation for what many science educators have long advocated: the integration of science practices and disciplinary ideas, with a focus on carefully chosen disciplinary core ideas that are generative and more opportunities for authentic experiences with doing science. In practice, this means learning is guided by clearly articulated, focused learning goals that anchor rich experiences with the science and engineering practices (SEPs) and crosscutting concepts (CCCs) and disciplinary core ideas (DCIs) together—and the integration of these three domains, called *three-dimensional learning*. In this chapter, we lay out criteria that represent the vision of what instructional materials should look like as the NGSS are more widely adopted. Our criteria are organized into four broad categories that we have articulated in the form of assertions about high-quality materials that support the vision of the NRC *Framework* and the NGSS:

1. Instructional materials support NGSS-driven learning goals.
2. Instructional materials provide coherence across the three dimensions.
3. Instructional materials support science learning experiences that promote three-dimensional learning.
4. Instructional materials provide ways to monitor student learning across the three dimensions.

Within each of these categories are evaluative criteria specifically focused on student materials, or those materials directly used by students in classrooms (e.g., textbooks, workbooks, handouts, accompanying videos or multimedia resources). There are also evaluative criteria for teacher materials to examine how well the accompanying teacher materials support teacher learning and implementation of materials (e.g., teacher editions of textbooks, pacing guides, additional multimedia or video resources). This reflects the finding that instructional materials need to be educative for teachers; that is, materials should “help to increase teachers’ knowledge in specific instances of instructional decision making, but also help them develop more general knowledge that they can apply flexibly in new situations” (Davis & Krajcik, 2005, p. 3).

Chapter 2 is organized so that the 11 criteria for student materials are parallel with the 11 criteria for teacher materials, and these criteria are each nested within one of the four broad categories (table 2.1).
Table 2.1. Criteria for evaluation of instructional materials in science.

<table>
<thead>
<tr>
<th>NGSS-Driven Learning Goals</th>
<th>Coherence across Three Dimensions</th>
<th>Learning Experiences across Three Dimensions</th>
<th>Monitoring Learning across Three Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1S</strong> Materials are based on learning goals. Those goals call for learning of</td>
<td><strong>4S</strong> Materials are designed with carefully sequenced learning goals and well-matched experiences.</td>
<td><strong>6S</strong> Materials provide multiple opportunities for students to share and negotiate their ideas, prior</td>
<td><strong>9S</strong> Materials include accessible and unbiased formative and summative assessments of students’ three-dimensional learning.</td>
</tr>
<tr>
<td>• disciplinary core ideas, science and engineering practices, crosscutting concepts</td>
<td><strong>4T</strong> Materials communicate the design principles and sequencing underpinning the storyline.</td>
<td>knowledge, and experiences.</td>
<td><strong>9T</strong> Materials highlight formative and summative assessments and provide tools and guidance for interpreting evidence of three-dimensional learning and using assessment results to plan for future instruction.</td>
</tr>
<tr>
<td>• the nature of science, engineering, technology, and applications of science from NGSS; and</td>
<td></td>
<td><strong>7S</strong> Materials use motivating contexts to engage students in real-world phenomena and authentic design</td>
<td><strong>10S</strong> Materials include multiple opportunities for self-assessment and reflection to promote sensemaking among students.</td>
</tr>
<tr>
<td>• Common Core State Standards for English language arts and mathematics.</td>
<td></td>
<td>problems.</td>
<td><strong>10T</strong> Materials provide guidance for teachers to use data from assessments to provide feedback to students and promote student self-assessment and reflection.</td>
</tr>
<tr>
<td><strong>1T</strong> Materials explain the learning goals; the rationale for selecting them; and</td>
<td></td>
<td><strong>5S</strong> Materials provide students with opportunities to make links across the three dimensions to build</td>
<td></td>
</tr>
<tr>
<td>• how they promote three-dimensional learning;</td>
<td></td>
<td>coherent conceptual understanding and abilities to use the practices.</td>
<td></td>
</tr>
<tr>
<td>• how they promote learning of the nature of science, engineering, technology, and</td>
<td></td>
<td><strong>5T</strong> Materials promote teacher knowledge-building related to the storyline.</td>
<td></td>
</tr>
<tr>
<td>• applications of science; and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• how they promote learning of the Common Core standards for English language arts and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mathematics.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2S</strong> Materials use phenomena or problems to focus students on learning goals.</td>
<td><strong>2T</strong> Materials explain how the phenomena or problems are used to focus students on learning goals.</td>
<td><strong>7T</strong> Materials provide guidance to teachers for using effective teaching strategies that engage students</td>
<td></td>
</tr>
<tr>
<td><strong>3S</strong> Materials are based on scientifically accurate and grade-level-appropriate</td>
<td><strong>3T</strong> Materials situate learning goals within the progression of K-12 learning laid out by the NGSS.</td>
<td>in real-world phenomena and authentic design problems.</td>
<td></td>
</tr>
<tr>
<td>learning goals.</td>
<td></td>
<td><strong>8S</strong> Materials are accessible to a wide range of students.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>8T</strong> Materials provide suggestions for how to address a range of students’ skills, needs, and interests.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When evaluating student materials, consider the extent to which …

When evaluating teacher materials, consider the extent to which …
Category A: Instructional Materials Support NGSS-Driven Learning Goals.

The NGSS are arguably the best resource to begin designing learning goals for instructional materials. While the standards are not necessarily learning goals in themselves, they help to determine grade-level-appropriate SEPs, CCCs, and DCIs that can serve as the basis for learning goals. The NGSS are written to capture increasingly sophisticated disciplinary knowledge and practice as students progress through K-12 education. In any given standard, however, there is a wealth of content or practices at play, and therefore, to design well-crafted learning goals, the standards must be “unpacked” into manageable learning goals.

Unpacking a standard is a process in which the standard is broken down into significant parts, with each part considered in light of the emerging capabilities of students at the intended age for instruction (Krajcik McNeill, & Reiser, 2008). Learning goals are then designed to focus on selected parts of the SEPs, CCCs, DCIs, and performance expectations, using real-world phenomena or design problems as the context for learning. The learning goals that emerge from the unpacking process may only partially address a standard or perhaps parts of several standards, but they typically do not cover the entirety of a standard as written.

Further, to achieve the vision of three-dimensional learning in the NRC Framework and the NGSS, the unpacking process must also involve integration of the relevant SEPs, CCCs, and DCIs in learning goals. Historically, however, instructional materials have separated content-based goals from practice- or skill-based goals. This divide between the content and practices of science not only misrepresents science itself, but may also be unproductive for science learning (Duschl, Schweingruber, & Shouse, 2007; Edelson, 2001). Well-constructed learning goals intertwine the NGSS dimensions at strategic points of instruction, presenting disciplinary knowledge accurately and accessibly, using relevant science and engineering practices, and emphasizing powerful crosscutting concepts.

Underscoring the emphasis on SEPs, CCCs, and DCIs is the need for science education to do a much better job at communicating a more authentic nature of science to students. Students often learn science as though it were a fixed body of knowledge, yet the actual nature of science is often exploratory, driven by questions. We know from many research studies that when science is treated as a body of factual knowledge, the concepts students learn tend to become “inert” and inaccessible (e.g., Bereiter & Scardamalia, 1985). Science is a social enterprise and a human endeavor, one that employs rigorous standards for scientific work but is ultimately dynamic, changeable, and open to revision in light of new evidence (NRC, 2012). This view of science can be taught to students explicitly through engaging with science and engineering practices—all practices that emphasize how scientific knowledge is constructed, validated, and revised over time.

The NGSS also highlights and encourages real-world engineering, technology, and applications of science, which help to communicate the relevance of science to students and provide authentic contexts in which to engage across the three dimensions. Designers of instructional materials need to think carefully about the phenomena and real-world contexts and applications they choose to use so that the scenarios and tasks support the learning goals while also being relevant to young people and attending to the practicalities of the science classroom. Finally, instructional materials can also support connections to other disciplines, such as literacy and mathematics, by helping students construct verbal and written explanations and arguments, analyze scientific texts, apply computational thinking to data sets, and participate productively in social interactions grounded in the science practices.

The first evaluative criterion within Category A focuses on how well the learning goals within instructional materials serve to bring about the vision of the NGSS, particularly three-dimensional learning, including the nature of science; engineering, technology, and applications of science; and connections to literacy and mathematics. Another hallmark of the Framework and the NGSS is the importance of focusing the learning goals around phenomena or design problems. This is discussed within the second criterion. Finally, situating
learning goals within the broader progression of learning is increasingly recognized as important for supporting and monitoring students’ progress over time (Corcoran, Mosher, & Rogat, 2009; NRC, 2012) and for building inter-unit coherence (Fortus, Sutherland Adams, Krajcik, & Reiser, 2015). This is discussed in the third criterion.

Because the evaluative criteria for this category set a high bar for the design of learning goals, teachers will need insight into the selection and purpose of these goals. The focus of the criteria for teacher materials is on how well the materials support the teachers in developing these insights. Learning goals written with this depth and complexity can be notably different from the learning goals currently familiar to teachers. Designers of instructional materials should recognize this departure from the familiar and be transparent about the design of the learning goals. This will support teachers’ ability to use the materials effectively as well as their capacity to adapt materials to meet students’ needs without losing sight of the main goal(s) (Davis & Krajcik, 2005; Davis & Varma, 2008; Penuel, Phillips, & Harris, 2014; Remillard, 2005).

The following evaluative criteria address the central role the NRC Framework and the NGSS should play in designing learning goals of materials. Student materials …

1S. are based on learning goals. Those goals call for learning of
- disciplinary core ideas, science and engineering practices, and crosscutting concepts from NGSS integrated as three-dimensional learning;
- the nature of science, engineering, technology, and applications of science from NGSS; and
- Common Core State Standards for English language arts and mathematics.

2S. use phenomena or design problems to focus students on learning goals.

3S. are based on scientifically accurate and grade-level-appropriate learning goals.

Teachers will need guidance in enacting materials with integrity while also making necessary accommodations. This requires supporting teachers as they learn to interpret and use new materials. We propose three criteria for teacher materials. Teacher materials …

1T. explain the learning goals; the rationale for selecting them; and
- how they promote three-dimensional learning;
- how they promote learning of the nature of science, engineering, technology, and applications of science; and
- how they promote learning of the Common Core State Standards for English language arts and mathematics.

2T. explain how the phenomena or problems are used to focus students on learning goals.

3T. situate learning goals within the progression of K-12 learning laid out by the NGSS.
Three-dimensional learning is exemplified by the NGSS Performance Expectations. The Performance Expectations represent a large-scale assessment framework and demonstrate how the three dimensions of NGSS might be integrated and ultimately realized in classrooms, but they are not an exhaustive list of the ways standards might be integrated given the ideas, practices, and concepts at a given age level (NGSS Lead States, 2013). But how are the Performance Expectations related to learning goals? In some cases, a Performance Expectation may align well with a unit’s learning goal, potentially serving this role.

In other cases, a Performance Expectation may encompass too many ideas, practices, or concepts, and, therefore, several learning goals would be necessary to scaffold students’ experiences toward the Performance Expectation. There is also the possibility that a different combination of SEPs, CCCs, and DCIs could

---

**When evaluating student materials, consider the extent to which ...**

**Evaluative Criterion 1S:** Materials are based on learning goals. Those goals call for the learning of

- disciplinary core ideas, science and engineering practices, and crosscutting concepts from NGSS integrated as three-dimensional learning;
- the nature of science, engineering, technology, and applications of science from NGSS; and
- Common Core State Standards for English language arts and mathematics.

**A Note about the Science and Engineering Practices**

Practices are not skills or tools, per se, but are much like the habits of mind or ways of knowing and engaging in the world. These habits of mind are acquired by scientists through repeated authentic use and are notably different from our everyday ways of thinking (Berland et al., 2015; O’Connor & Michaels, 1996; Resnick, 1987; Rosebury, Warren, & Conant, 1992). Argumentation in science, for example, is based on logical reasoning and evidence, whereas the argumentation we experience in our daily lives might be based less on logic and evidence and more on our emotions and cultural background (Bricker & Bell, 2008; Moje et al., 2004). Importantly, learning goals need to focus instructional activities on the repeated authentic use of practices, including experiences in which teachers model practices for students and scaffold students’ use of practices over time (Krajcik, McNeill, & Reiser, 2008). Instructional material designers cannot assume that students will readily engage in science practices without sufficient support that includes multiple and varied opportunities to engage in practices, especially in social learning situations.
produce a more appropriate learning goal for a unit. Regardless, the emphasis of the learning goal should be on integrating the three dimensions during the instructional sequence; that is, three-dimensional learning may not be present at all times but should be a driving force for how students engage across an instructional sequence. Consider some examples of how materials might integrate only one dimension, two dimensions, or all three dimensions (dimensions highlighted using color key below):

- **One dimension**: Students describe **climate** as a typical range of weather conditions in an area and the extent to which those conditions vary over years.

- **Two dimensions**: Students analyze and interpret maps of average temperature and rainfall amounts to determine the typical range of weather conditions in an area and over time.

- **Three dimensions**: Students analyze and interpret maps of average temperature and rainfall amounts to identify patterns in the typical range of weather conditions in a variety of areas and over time.

While learning goals can strategically integrate the three dimensions of the NGSS to promote student learning, they can also provide students with a more authentic experience with the dynamic nature of science, engineering, technology, and the applications of science. The NGSS propose a radical shift from students learning only about what scientists know to both what scientists know and how scientists and engineers think and engage in their work. The nature of science at its very essence is about exploring curiosities, asking questions, and investigating, all in the name of explaining how the world works. Applied sciences and engineering also seek design-based solutions that manipulate and alter the natural world to solve problems or open new possibilities. Given the current state of instructional materials and the way science is presented to students, most students believe science knowledge to be permanent, objective, and logical in nature as opposed to tentative, partially subjective, and often creative (Zeidler, Walker, Ackett, & Simmons, 2002). Research has also shown that students can transform their thinking about the nature of science through experiencing intentional, explicit instruction and not just implicitly by engaging with science and engineering practices (Khishfe & Abd-El-Khalick, 2002; Lederman, 2007).

The student experience with the nature of science, engineering, technology, and applications of science can also be enhanced by the explicit connection in the NGSS to the Common Core State Standards for English language arts and mathematics, as specified with the NGSS. Understanding the social enterprise of science also draws upon practices typically found in other disciplines, such as argumentation taught in language, speech, or debate classes or the use of primary source documents used frequently in the social sciences. The abilities to engage in complex mathematical reasoning and to be critical readers and producers of text play a large role in scientific work (Bazerman, 1988). Likewise, scientifically literate citizens understand the tentative nature of science, which is based on the best evidence available at the time, and they know that science is a social enterprise with rigorous standards for acceptance by the larger science community but with room for human error.
Instructional materials play a central role in supporting connections to language and mathematics (see table 2.2). Research has shown that instructional programs that emphasize strategic reading in the content areas improve students’ comprehension of concepts and their engagement with the content (e.g., Guthrie et al., 2004; Pearson, Moje, & Greenleaf, 2010). Writing in science also improves student understanding of science and increases students’ awareness and ownership of what is being learned (Prain & Hand, 1999). The act of engaging in verbal and written explanations and argumentation can help to build students’ language and literacy skills while also engaging students in the social validation process that is a truer reflection of the nature of science. Like language arts, mathematics has an important role to play in science instruction. Students will be expected to collect, organize, analyze, and interpret data in order to construct explanations and arguments about a phenomenon or design problem.

Table 2.2. Examples of connections to Common Core State Standards excerpted from Appendix L and M in NGSS Volume 2.

<table>
<thead>
<tr>
<th>Connection to Common Core Standard</th>
<th>Science example (Appendix L) or connection to Science and Engineering Practice (Appendix M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5ESS1 Earth’s Place in the Universe with 5.G.A.2: Represent real world and mathematical problems by graphing points in the first quadrant of the coordinate plane, and interpret coordinate values of points in the context of the situation.</td>
<td>Science examples from 5ESS1 Earth’s Place in the Universe: (1) Over the course of the year, students compile data for the length of the day over the course of the year. What pattern is observed when the data are graphed on a coordinate plan? How can a model of the sun and Earth explain the pattern? (2) Students are given ((x, y)) coordinates for the Earth at six equally spaced time during its orbit around the sun (with the sun at the origin). Students graph the points to show snapshots of Earth’s motion through space. Example from p. 148 of NGSS Volume 2 Appendix L.</td>
</tr>
<tr>
<td>CCR Writing Anchor #7: Conduct short as well as more sustained research project based on focused questions, demonstrating understanding of the subject matter under investigation.</td>
<td>Planning and carrying out investigations to test hypotheses or designs is central to scientific and engineering activity. The research practices reflected in Writing Standard 7 reflect the skills needed for successful completion of such research-based inquiries. Example from p. 161 of NGSS Volume 2 Appendix M.</td>
</tr>
</tbody>
</table>

Supporting students in learning across the three dimensions of the NGSS will require a transformational shift in the way the nature of science is communicated in schools. This shift will naturally entail an explicit focus on the nature of science, with connections to engineering, technology, and applications of science, language and literacy, and mathematics as natural extensions of how scientists conduct their work.
### Recommended indicators for evaluating this criterion

1. Are elements of the NGSS three dimensions evident within the learning goals?
   - a. Identify at least one science and engineering practice that is a focus of the unit. Specifically, which element* of the practice is evident? How well do the learning goals promote sustained engagement and progressively more sophisticated experience with the practice?
   - b. Identify at least one crosscutting concept that is the focus of the unit. Specifically, which element* of the concept is evident? How well do the learning goals promote sustained engagement and progressively more sophisticated experience with the concept?
   - c. Identify at least one disciplinary core idea that is the focus of the unit. Specifically, which element* of the disciplinary core idea is evident? How well do the learning goals promote sustained engagement and progressively more sophisticated experience with the core idea?

2. How well do the learning goal(s) integrate SEPs, CCCs, and DCIs?

3. If instruction is guided by multiple learning goals, what is the extent to which these are an appropriate combination of goals that contribute toward three-dimensional learning?

4. Within a lesson, instructional sequence, chapter, or unit:
   - a. Is the nature of science highlighted for students where appropriate? (NGSS, Appendix H, 2013)
   - b. Is engineering, technology, and applications of science highlighted for students where appropriate? (NGSS, Appendices I & J, 2013)
   - c. Are reading, writing, speaking, and/or listening (Common Core ELA) incorporated in meaningful and productive ways?
   - d. Are mathematical or computational concepts or practices (Common Core Mathematics) incorporated in meaningful and productive ways?

5. To what extent are the connections appropriate within the activity or learning sequence? Do they seem to be a natural fit?

6. To what extent are the connections made explicit to students?

* Elements are the bulleted sub-points of the dimension represented in the foundation boxes of the standards and in K-12 grade-banded progressions in the Appendices E, F, and G.
Learning goals focus students on the science phenomena or design problems they will be studying over an instructional sequence. Phenomena and engineering design problems are central to the NRC Framework and the NGSS, as “the goal of science is to develop a set of coherent and mutually consistent theoretical descriptions of the world that can provide explanations over a wide range of phenomena” (NRC, 2012, p. 48). Learning goals can also use phenomena or design problems to make student engagement purposeful (e.g., Duschl, 2008; Krajcik, McNeill, & Reiser, 2008; Marzano, 2007; Roth et al., 2011; Zembal-Saul, McNeill, & Hershberger, 2013). Providing students with a sense of purpose, especially during hands-on investigations, can significantly improve student learning (Barron et al., 1998; Blumenfeld et al., 1991). This is especially true when students are expected to engage with science practices that may be unfamiliar to them (Krajcik, Slotta, McNeill, & Reiser, 2008).

To use phenomena or design problems to focus students on learning goals, materials might begin an instructional sequence with a guiding (or driving) question, a discrepant event, or puzzling phenomena to engage students in the instruction. Selecting a question, discrepant event, or phenomenon that aligns closely with the learning goal will engage students, give them a sense of purpose, and provide a more interesting entry into the instructional sequence. Posing a puzzling question—for example, “Is a cloud air or water?”—is more likely to pique students’ curiosity than a well-written learning goal, such as, “Students will be able to provide evidence that a cloud is a mass of condensed water vapor floating in the atmosphere.” Likewise, an instructional unit might initiate learning with a discrepant event (e.g., dropping a heavy ball and a light ball at the same time and asking students to predict which will land first) or a design problem (e.g., choosing appropriate materials based on their characteristic properties to create a solution to a problem). Situating the learning using a puzzling phenomenon, interesting question, discrepant event, or design problem can make the learning goal more appealing and motivating to students.

Whatever approach is used to introduce the instructional sequence to students, the phenomenon or design problem driving the instructional sequence should be clear to students and well connected to the learning goal(s), and, at some point in the sequence, the learning goals should be made explicit to students in a language that is appropriate and understandable to them. The learning goals must be stated in a complete sentence or question with a level of detail sufficient to clearly understand the expectations for student learning, as opposed to listing topics of study that might be too broad or vague to convey the goal for student learning (see figure 2.1 for examples).
Clear* Learning Goals or Guiding Questions to Explain Phenomena

- Learning Goal: Interpret maps to explain how water changes the shape of the land through erosion and deposition.
  - Guiding Question: Can mountains grow so tall they can reach outer space?
- Learning Goal: Construct an explanation of the role of photosynthesis in the cycling of matter and flow of energy into and out of organisms.
  - Guiding Question: How do plants make food?
- Learning Goal: Use evidence to identify which chemical reactions release energy or store energy and then hypothesize why there is a difference between the two.
  - Guiding Question: What makes an ice pack cold? What makes a hot pack hot?

* Clear, explicit learning goals for students are stated in complete sentences with enough detail that students (and teachers) understand the expectations for learning (Roth et al., 2011).

Figure 2.1. Examples of clear learning goals based on explaining phenomena.

Recommended indicators for evaluating this criterion

1. Are learning goals within the lesson, instructional sequence, chapter, and/or unit focused on explaining natural phenomena or designing solutions to problems?
   a. Is the phenomenon or design problem appropriately connected to the learning goal(s)?
   b. Do the learning goals (or guiding questions) communicate a clear sense of purpose for student engagement in the instructional activities?
2. Is the use of phenomena or design problems productive for advancing students’ understanding of the three NGSS dimensions?
3. Is the language of learning goals (or guiding questions) accessible to students?
   a. Are learning goals worded in a way that students can understand?
   b. Do learning goals avoid unnecessary scientific jargon that can confuse students?

When evaluating student materials, consider the extent to which ...

Evaluative Criterion 3S: Materials are based on scientifically accurate and grade-level-appropriate learning goals.

One of the conceptual shifts in the NGSS is that science concepts build coherently through K-12, providing a connected and increasingly sophisticated learning of science across grade bands. The NGSS states, “What this means to teachers and curriculum developers is that the same ideas or details are not covered each
year. Rather, a progression of knowledge occurs from grade band to grade band that gives students the opportunity to learn more complex material, leading to an overall understanding of science by the end of high school” (NGSS Lead States, 2013, Appendix A, p. 3).

Regardless of whether districts or states have adopted the NGSS or developed different standards based on the Framework, these standards provide the most up-to-date guidance on the appropriate grade level or grade band in which science concepts and practices should be taught. The elements of the NGSS dimensions have been placed in grades or grade bands based on what we know from research on student learning in science. One line of research that has been particularly influential is learning progressions, which describe how science ideas and/or practices deepen across grade levels and grade bands (Corcoran et al., 2009; Mosher, 2011; NRC, 2012). This relatively new body of research builds on decades of cross-sectional research on common student conceptions, which describe student capabilities and developmental trajectories at given ages, mostly in terms of concept development. From this research we know that students in the early years of education have partially formed ideas about phenomena, with more or less accuracy. These partially formed ideas can act as productive “stepping stones” toward more sophisticated learning later on (Wiser, Smith, & Doubler, 2012), or they can prevent students from mastering more sophisticated content if they are not addressed.

Appendices E, F, and G in the NGSS show how SEPs, CCCs, and DCIs deepen across grade levels, allowing developers of instructional materials to pinpoint appropriate learning at one grade level or grade band within the larger landscape of instruction. As developers of instructional materials unpack standards to design learning goals, they can use these resources to ensure that the learning expected from students is appropriate for the intended age and is productive for richer learning experiences in subsequent grade levels (NRC, 2012). Similarly, evaluators can determine the accuracy and grade-level appropriateness of the learning goals with materials by using the same resources.

**Recommended indicators for evaluating this criterion**

1. Are the learning goals in the student instructional materials grade-level appropriate?
   a. Do the learning goals fit within the targeted grade band and the larger progression of learning across grade bands as stated by the NGSS?
   b. Given what students have learned previously, does the learning goal appropriately challenge students?
   c. Does the learning goal adequately position students for deeper learning of the NGSS dimensions later on?

2. Is the science content within the learning goals scientifically accurate?
While specifying learning goals within teacher materials is commonplace (Dick, Carey, & Carey, 2015), often teachers are not given the rationale behind the goals. When learning goals are not explicit and/or the rationale for the goals is not provided to teachers, teachers may struggle with staying focused on the goal(s) in general or when they make adaptations to the materials and instruction for their particular students and contexts. In such situations, teachers may default to a familiar approach to instruction, which often varies greatly from the developers’ intentions (Reiser, Krajcik, Moje, & Marx, 2003). One way to improve the integrity of implementation of the materials is to make the learning goals and rationale for their selection transparent and to explain how certain activities support these goals. This information can help teachers make informed adaptations in their classrooms while also staying focused on the main learning goals.

Instructional materials can also provide guidance to support teachers in successfully using threedimensional learning goals to guide their classroom instruction. Because three-dimensional learning is a significant shift in the science standards, instructional materials will need to be explicit to teachers about how learning goals build toward three-dimensional learning during an instructional sequence. This will be especially important as teachers guide students in learning experiences (e.g., Krajcik, McNeill, & Reiser, 2008; Quintana et al., 2004). Given the pre-NGSS approach of often separating content goals from skill or practice goals, instructional materials that exemplify the NGSS will need to provide guidance to teachers about how to scaffold three-dimensional learning over time in pursuit of an integrated learning goal, starting with an appropriate entry point at the beginning of an instructional unit and working toward more challenging, three-dimensional tasks as the unit unfolds.

Teacher materials can support teachers in understanding the connection between instructional activities and three-dimensional learning goals by highlighting how activities support the learning goals. For example, materials can highlight three-dimensional learning using icons to indicate when specific SEPs, CCCs, and DCIs are integrated in the experience. Teacher materials can also provide teachers with videos of classroom learning experiences that embody three-dimensional learning or artifacts of student work that represent three-dimensional learning.

When evaluating teacher materials, consider the extent to which …

Evaluative Criterion 1T: Materials explain the learning goals, the rationale for selecting them, and

• how they promote three-dimensional learning;
• how they promote learning of the nature of science, engineering, technology, and applications of science; and
• how they promote learning of the Common Core State Standards for English language arts and mathematics.

Two ideas frame this criterion—the importance of teachers’ understanding of intended learning goals and their ability to adapt materials to meet the needs of students. Adapting instructional materials to meet student and contextual needs is a hallmark of an excellent teacher. There is no one set of materials that works equally well with all students; thus, an exceptional teacher will adapt materials to better reach all learners in his or her classroom. However, teacher adaptations can either enhance learning of the intended goals or lead learners away from the learning goals. Classroom research shows that when teachers are familiar with the theoretical underpinning of learning goals and the ultimate intentions of the materials, teachers can enact curriculum with integrity, retaining the core essence of the materials while also adapting them to their context (Connelly & Ben-Peretz, 1997; Davis & Varma, 2008). Additionally, when teachers are clear about the learning goals, students are more likely to achieve them (e.g., Barron et al., 1998).

While specifying learning goals within teacher materials is commonplace (Dick, Carey, & Carey, 2015), often teachers are not given the rationale behind the goals. When learning goals are not explicit and/or the rationale for the goals is not provided to teachers, teachers may struggle with staying focused on the goal(s) in general or when they make adaptations to the materials and instruction for their particular students and contexts. In such situations, teachers may default to a familiar approach to instruction, which often varies greatly from the developers’ intentions (Reiser, Krajcik, Moje, & Marx, 2003). One way to improve the integrity of implementation of the materials is to make the learning goals and rationale for their selection transparent and to explain how certain activities support these goals. This information can help teachers make informed adaptations in their classrooms while also staying focused on the main learning goals.

Instructional materials can also provide guidance to support teachers in successfully using threedimensional learning goals to guide their classroom instruction. Because three-dimensional learning is a significant shift in the science standards, instructional materials will need to be explicit to teachers about how learning goals build toward three-dimensional learning during an instructional sequence. This will be especially important as teachers guide students in learning experiences (e.g., Krajcik, McNeill, & Reiser, 2008; Quintana et al., 2004). Given the pre-NGSS approach of often separating content goals from skill or practice goals, instructional materials that exemplify the NGSS will need to provide guidance to teachers about how to scaffold three-dimensional learning over time in pursuit of an integrated learning goal, starting with an appropriate entry point at the beginning of an instructional unit and working toward more challenging, three-dimensional tasks as the unit unfolds.

Teacher materials can support teachers in understanding the connection between instructional activities and three-dimensional learning goals by highlighting how activities support the learning goals. For example, materials can highlight three-dimensional learning using icons to indicate when specific SEPs, CCCs, and DCIs are integrated in the experience. Teacher materials can also provide teachers with videos of classroom learning experiences that embody three-dimensional learning or artifacts of student work that represent three-dimensional learning.
Additionally, teachers will need guidance on how to support the nature of science and connections to engineering, technology, and the applications of science. Materials can assist teachers by presenting science as both a logical investigative process and a process of social construction through talk, text, and community validation. Students need to read, examine, and discuss primary source documents created during the time of scientific revolutions and about how scientists debated and disagreed over what we now take as common knowledge. Students need to read and discuss present-day science and engineering revolutions and witness the social aspects of the scientific process. Instructional materials will need to provide guidance to teachers on how to approach these topics in class so that students see not only that science is debatable and changeable but also that science follows rigorous, strict standards. That is, our explanations only change with sufficient evidence and community validation of that evidence. Because many teachers have never themselves had opportunities to learn about these aspects of science, they too may need opportunities to learn more about the scientific and engineering enterprise.

Finally, English language arts and mathematics are a focal point in every child’s education. They are not exclusive of science, but rather, learning science can be enhanced by reading, writing, and mathematics integration (NGSS Lead States, 2013). Including the Common Core State Standards for English language arts and mathematics into science instruction can provide a comprehensive education for students that mirrors much more authentic science and engineering disciplines beyond K-12 education. For example, analyzing informational text, writing an explanation based on evidence, engaging in argumentation and discourse with peers, and using mathematics to understand scale, proportion, and quantity are just a few examples of how science requires mastery of mathematics and English language arts. Science teachers, however, are not mathematics teachers nor are they English language arts teachers. While elementary school teachers and some secondary school teachers may feel comfortable with integrating mathematics and literacy in science instruction, many science teachers will need assistance in doing this well. Designers of instructional materials need to identify a set of Common Core standards that connect well to the science practices and concepts in the materials (see NGSS Appendices L and M) and then provide strategies to incorporate reading, writing, speaking, and listening as well as mathematics strategies across instructional activities.

Considerations for evaluating this criterion should include how well the teacher materials highlight three-dimensional learning, the nature of science and applications of science, and the connections to English language arts and mathematics in relation to the learning goal(s). Materials should be evaluated by how well they help teachers see when and how activities or text contribute to the learning goals across instructional sequences or book chapters and supplemental resources.
Recommended indicators for evaluating this criterion

1. Is the learning goal(s) explicitly stated at the beginning of an instructional sequence, chapter, or unit within the teacher materials?

2. Is the learning goal(s) highlighted for teachers and clearly worded so that teachers understand the goal(s) for instruction?

3. Do the teacher materials provide insights about the intentions and purpose of the learning goal(s)?

4. Do the teacher materials highlight when and how to make learning goals explicit throughout the instructional sequence, chapter, or unit?
   a. Do the materials point to key instructional activities that support the overall development of the learning goal(s)?

5. Do the materials suggest when and how adaptations could be made without detracting from the learning goal(s)?

6. Do teacher materials clearly articulate and highlight the three dimensions of learning within and across instructional sequences as they support the learning goal(s)?

7. Do the teacher materials provide information about how to scaffold student engagement with the three dimensions across time toward the learning goal(s)?
   a. Of particular importance, do the teacher materials provide information about scaffolding engagement in science and engineering practices and crosscutting concepts so that these are not forgotten in pursuit of disciplinary core ideas?

8. How well are the following connections highlighted for teachers? How well do materials support teachers in making the following connections or in helping students make the connections?
   a. Are appropriate connections to the nature of science highlighted and explained for teachers (NGSS, Appendix H, 2013)? Do the materials support teachers in explicitly talking about and engaging with the nature of science?
   b. Are appropriate connections to engineering, technology, and applications of science highlighted and explained for teachers (NGSS, Appendices I & J, 2013)? Are these connections a cornerstone or an afterthought?
   c. Are reading, writing, speaking, and/or listening (Common Core ELA) highlighted for teachers with support for implementation? Do the materials make it clear to teachers how oral and written student work supports both science and literacy goals?
   d. Are mathematical or computational concepts or practices (Common Core Mathematics) highlighted for teachers with support for implementation? Do the materials make it clear to teachers how math/computational/graphing activities support both science and math goals?
   e. Are other relevant disciplinary concepts or practices, such as those from the social sciences, highlighted for teachers with support for implementation?
The NRC *Framework* and the NGSS have posed a shift away from students learning science ideas disconnected from explaining real-world phenomena or designing solutions to engineering problems (e.g., NRC, 2015, p. 11). It will no longer be sufficient for students to simply memorize science definitions when learning goals are based on explaining phenomena or designing solutions to problems. Rather, students will be expected to use science ideas to generate explanations and solutions (Reiser, 2013). For example, in some current instructional materials, students are asked to identify the arrangement of water molecules in different states of matter—separate from any connection to other science ideas. In a phenomena-based instructional sequence, the same science ideas about water at the molecular level can be used to help students explain weather phenomena. While the learning goals might be focused on states of matter and phase changes in water, students could be asked to use those concepts to explain how clouds form.

This focus on explaining phenomena and designing solutions represents another significant shift in science learning for both students and teachers. Teacher materials will need to support teachers in making this shift by explaining how phenomena and engineering design problems are used as a focus for instruction. At the beginning of an instructional sequence, teacher materials can provide a clear connection between the science ideas and how they can be used to explain a chosen phenomenon or to design a solution to a problem. Throughout the instructional sequence, teachers can also be given guidance to revisit the phenomenon or problem to help students piece together and revise an explanation or a solution as they develop new understandings about science ideas.

**When evaluating teacher materials, consider the extent to which ...**

**Evaluative Criterion 2T: Materials explain how the phenomena or problems are used to focus students on learning goals.**

---

**Recommended indicators for evaluating this criterion**

1. Do the teacher materials explicitly state the phenomenon or engineering design problem that is the focus of the instructional sequence?

2. Do the teacher materials make clear connections between the science ideas in the instructional sequence and how they can help students explain the phenomenon or design solutions to engineering problems?

3. Are there points in the instructional sequence where teachers are provided guidance to revisit the phenomenon or design problem with students to develop and revise explanations and solutions?

---

**When evaluating teacher materials, consider the extent to which ...**

**Evaluative Criterion 3T: Materials situate learning goals within the progression of K-12 learning laid out by the NGSS.**
Most science educators are familiar with the idea of a developmental progression of learning across time. In practice, developmental progressions often occur as “scope and sequence” documents that curriculum specialists and teachers consult to understand the target ideas at specific grade levels and to delineate boundaries with neighboring grade levels. Looking beyond a single year of instruction is necessary to see how disciplinary knowledge and practice will grow for a student over time. Most scope and sequence documents are constructed and validated using logic from the expert’s perspective; that is, they are not necessarily grounded in the emerging capabilities of students identified by research on learning (Corcoran et al., 2009). Like scope and sequences, the thoughtful and detailed Atlas maps produced by AAAS Project 2061 (AAAS, 2001, 2007) outline a logical progression of conceptual targets at different grade levels building toward increasingly more sophisticated understanding. The Atlas maps are especially useful for seeing the interconnectedness of science ideas across time. However, the growing body of research on student thinking, especially from research on learning progressions, has shown the developmental progression of understanding to be even messier than the Atlas maps would suggest (Gotwals & Songer, 2010).

School districts and teachers will continue to need organizing tools, like the Atlas maps, for their curriculum and can readily use them when provided. Given the growing contributions of learning progressions research to our understanding of how students learn, developers of instructional materials will need to rethink how to present the scope and sequence to practitioners, considering where progressions grounded in research on student learning support and deviate from the current top-down progressions.

The scope and sequence of materials that exemplify the NGSS should be based on the current, research-based understandings of learning progressions where that research exists. If designers include such organizing tools in teacher materials, what might result is something similar to Appendixes E, F, and G in the NGSS that shows clear boundaries between grade bands for the development of SEPs, CCCs, and DCIs. These documents, however, do not provide the level of detail necessary for a practitioner to see the starting point and end point for their students in a given school year. For example, when and how might teachers introduce modeling to their students and then scaffold engagement with modeling across the school year? Instructional materials are well positioned to offer tools and instructional supports to fill in the gaps from scope and sequence documents for teachers so they better understand their students’ likely entry points with the learning goals and reasonable attainment of those goals within the larger context of learning for students across their K-12 education.

**Recommended indicators for evaluating this criterion**

1. Do the materials include an organizing conceptual resource that shows where the elements of the SEPs, CCCs, and DCIs are situated within the broader curriculum landscape?

2. Do the teacher materials explain how the relevant disciplinary core ideas, science and engineering practices, and crosscutting concepts increase in rigor and sophistication across instructional sequences?
   
   a. Do the materials provide clear and explicit starting points for engaging with SEPs, CCCs, and DCIs?

   b. Do the materials provide scaffolds for increasing engagement with the SEPs, CCCs, and DCIs?

   c. Do the materials clearly identify the end points where students need to be after instruction for all three dimensions and how this relates to subsequent learning?
Category B: Instructional Materials Provide Coherence across the Three Dimensions.

Curriculum coherence has been posited as the single most important factor influencing student learning (Schmidt, Wang, & McKnight, 2005). Curriculum design research has shown that even though well-designed instructional materials that focus on coherence require more instructional time, students experience significant gains in learning (Roseman et al., 2013). Prior research also indicated that coherent science instruction favoring depth and rich knowledge integration paid off for student learning, but as designers streamlined units to reduce the instructional time required to complete the units, student learning diminished as well (Clark & Linn, 2003).

A key criticism of the US science education system has been the “mile-wide, inch deep” curricula that is pervasive in our schools (Schmidt, McKnight, Cogan, Jakwerth, & Houang, 1999; Schmidt, Wang, & McKnight, 2005). Science instructional programs tend to include far more content than can be covered in depth during a school year (Kesidou & Roseman, 2002; Roseman, Stern, & Koppal, 2010; Stern & Roseman, 2004), resulting in an overburdened and often incoherent curriculum (Schmidt, McKnight, & Raizen, 1997). They have been criticized for failing to make the interconnectedness of science ideas explicit to students (Roseman, Linn, & Koppal, 2008). This has led to tensions when science teachers are asked to teach concepts and practices in depth but are provided with instructional resources designed for breadth of coverage (Krajcik, McNeill, & Reiser, 2008). The NGSS emphasizes depth over breadth, but teachers still need to address many SEPs, CCCs, and DCIs and will need guidance in doing so productively in the limited instructional time available.

While the NGSS specifies the grade or grade band in which SEPs, CCCs, and DCIs should be taught, developers of instructional materials must determine when and how these components will be addressed. This work includes determining not only the sequence of learning goals and instructional activities within a single instructional unit but also the sequence of learning experiences across a school year and a grade band to achieve inter-unit coherence. Achieving coherence within and between units is difficult, and when it does not occur it often results in students learning science as discrete bits and pieces of information. This is especially true for instructional programs that place closely interrelated concepts in separate chapters (Roseman et al., 2008). While curricular coherence is currently hard to find in most science programs, knowing that it pays off for student learning provides the foundation for including it as an important evaluative category for instructional materials.

To this end, it is imperative that instructional materials offer coherent and focused instructional sequences and content to ensure that the three dimensions of the NGSS can be addressed in depth despite the practical limits on instructional time. Every activity in an instructional sequence (and across multiple sequences) must play a specific and necessary role in the overall science content storyline (Roth et al., 2011). When each instructional activity is closely linked to the main learning goal and content storyline and when students revisit science ideas and practices before, during, and after instruction, students show significant gains in learning (Roth et al., 2011; Taylor, Roth, Wilson, Stuhlsatz, & Tipton, 2016).

Evaluative criteria for student materials focus on the sequence of learning goals across chapter and units and how well the text, the learning experiences, and the performance tasks are matched to these goals. Materials must also be evaluated in terms of the connections made between science ideas and practices within and across individual lessons and units, especially in light of the shift to three-dimensional learning. The evaluation of materials must also consider how these ideas and practices engage students in sensemaking to develop conceptual understanding and in developing abilities to effectively use the practices independently. The following criteria should be used to evaluate the extent to which student
To achieve curricular coherence, we argue for materials that utilize a science content storyline. BSCS uses the following description of storyline:

*A storyline consists of carefully chosen and sequenced science ideas that build on one another to illustrate a bigger picture. This coherent set of science ideas creates a “story” within a lesson, as well as across lessons and units. The ideas flow from one to the next so that students can make the connections, just like they can follow and make sense of a good story. The central ideas of the story are emphasized, connected, and linked. Details are used to support the development of the central storyline, but are kept to a minimum so they don’t clutter and detract from the storyline.* (BSCS, 2015, p. 39)

Whether curriculum developers seek to create textbooks or project-based inquiry units that are coherent, the use of a science content storyline can help them determine how to sequence learning goals and develop supporting activities and text as the story unfolds. If an instructional sequence does not read like a logical narrative, then the storyline will likely not be visible to students either without substantial adaptation by the teacher. Just as any good narrative pieces together characters, settings, and plots to tell a story, the science content storyline should piece together the SEPs, CCCs, and DCIs, along with connections to the nature of science, to help students build toward meaningful understanding.

Coherent materials attend to the hierarchy of concepts as well as their sequence. The K-12 Alliance at WestEd developed both a product and process—conceptual flow—to represent these relationships (DiRanna, 1989; DiRanna, & Topps, 2004; DiRanna et al., 2008). In a conceptual flow, concepts are nested to show bigger ideas and supporting ideas. Concepts are sequenced to further clarify the relationships among the ideas. In light of the NGSS, practices and their elements need to be nested and sequenced over time to ensure coherence.
Within a school year and across grade levels, coherent learning experiences in science can help students develop a sophisticated network of science ideas. Without this coherence, discrete facts and bits and pieces of knowledge get mixed up, are difficult to retrieve, or get lost altogether. Studies have shown that the rich knowledge of experts in a field differs greatly from the disconnected knowledge of novices (e.g., Hmelo-Silver, Marathe, & Liu, 2007). Expert knowledge is highly interconnected, easy to retrieve, and transferrable to new contexts. Jerome Bruner argued that knowledge integration is key for helping students truly learn the wealth of information they are presented with in schools every day (Bruner, 1960, 1996).

The order in which science ideas, practices, and concepts are introduced should be carefully planned so that the sequence is clear for students and the flow of activities supports the development of the three dimensions across time. In addition to well-planned sequencing, the activities must closely match the learning goals so that students are able to make connections between them (Roth et al., 2011). Thus, each component of the instructional materials—from activity to lesson to chapter to unit—plays a specific role in the content storyline, with distracting activities and unnecessary details avoided (Harp & Mayer, 1998). Further, the sequence must make conceptual and developmental sense for students, which may differ from the way experts make conceptual sense of the content (Duncan & Rivet, 2013; Roseman et al., 2008). When a content storyline is well constructed, students show significant gains in learning as a result (Roseman et al., 2013; Roth et al., 2011).

### Recommended indicators for evaluating this criterion

1. Is the main learning goal (or set of learning goals) a central feature throughout the instructional sequence?
   a. Is there a clear main learning goal(s)?
   b. Is the main learning goal well matched to the content storyline?
   c. Does the sequence avoid distracting activities, terms, ideas, or text?

2. Does each activity focus on the main learning goal(s) and is it linked to appropriate SEPs, CCCs, and DCIs?

3. Do the learning experiences support productive and well-sequenced engagement with SEPs, CCCs, and DCIs?
   a. How well do activities within lessons enhance development of SEPs, CCCs, and DCIs?
   b. How well do activities across lessons enhance development of SEPs, CCCs, and DCIs?

4. Does the content storyline support increasingly more challenging engagement with SEPs, CCCs, and DCIs?

When evaluating student materials, consider the extent to which ...

**Evaluative Criterion 5S: Materials provide students with opportunities to make links across the three dimensions to build coherent conceptual understanding and abilities to use the practices.**

Within a school year and across grade levels, coherent learning experiences in science can help students develop a sophisticated network of science ideas. Without this coherence, discrete facts and bits and pieces of knowledge get mixed up, are difficult to retrieve, or get lost altogether. Studies have shown that the rich knowledge of experts in a field differs greatly from the disconnected knowledge of novices (e.g., Hmelo-Silver, Marathe, & Liu, 2007). Expert knowledge is highly interconnected, easy to retrieve, and transferrable to new contexts. Jerome Bruner argued that knowledge integration is key for helping students truly learn the wealth of information they are presented with in schools every day (Bruner, 1960, 1996).
Sadly, researchers have identified instructional materials that include extensive breadth of content with little or no connections between ideas and concepts as one of the key factors preventing students from building interconnected knowledge in science (Kesidou & Roseman, 2002; Roseman et al., 2008; Schmidt, McKnight, Valverde, Houang, & Wiley, 1997; Stern & Roseman, 2004). It is fairly common for US students to experience science as a set of discrete facts and activities with a weak connection to a larger conceptual understanding (Roth et al., 2006). Thus, the underlying science principles and their explanatory power are left to the students to discover without much support from instructional materials (Banilower, Boyd, Pasley, & Weiss, 2006; Roth & Garnier, 2006).

High-quality instructional materials can support sensemaking, conceptual understanding, and the development of science practices in several ways, such as using learning goals to drive the design of materials, anchoring instructional activities to learning goals, linking science ideas to each other, linking science ideas to instructional activities, and encouraging students to engage in sensemaking strategies (Ainsworth & Burcham, 2007; Bagno & Eylon, 1997; Clark & Linn, 2003; Krajick, McNeill, & Reiser, 2008; Linn, Clark, & Slotta, 2002; Roseman et al., 2013; Roth et al., 2011). Research shows that students who experience these types of instructional materials learn significantly more science than students who do not. For example, student learning improved for low and high achievers, female and male students, and students of different racial and socioeconomic backgrounds when instructional materials integrated the SEPs, CCCs, and DCIs (e.g., Harris et al., 2014).

### Recommended indicators for evaluating this criterion

1. Are the learning experiences sequenced in a conceptual and developmental way that will likely make sense to students?
   a. How well do the learning experiences within and across instructional sequences help to create a clear and meaningful flow of ideas for students?
   b. Does the sequence seem logical for students?

2. Do the materials include opportunities for students to revisit their learning around the SEPs, CCCs, and DCIs?
   a. Do the materials include strategies or experiences where students draw upon their existing knowledge?
   b. Do the materials include opportunities to elaborate or revise existing knowledge after an encounter with new information or experiences?
   c. Do the materials include opportunities for students to summarize and make connections back to the main learning goal(s)?

3. Do the materials make connections among ideas, activities, and science practices clear to students?
   a. Are SEPs, CCCs, and DCIs linked across instructional activities and are these links made explicit to students?
   b. Are science ideas linked to other ideas and are these links made explicit to students?
   c. Are science ideas linked to activities and are these links made explicit to students?
Most science teachers, especially at the secondary level, are familiar with the fast-paced, topical approach to teaching science content, focused on discrete chunks of information that are aligned to a standard and contained within a single chapter of a textbook. Teachers often supplement their textbook or other curriculum program to help their specific students master a standard. However, often teachers address one or two standards at a time, making progress through a textbook one chapter at a time. The district pacing guide becomes an essential reference document determining exactly what to teach and for how long, and teachers may eliminate activities or readings that do not align exactly with what their pacing guide dictates. Thus, if teachers are presented with NGSS-aligned materials that integrate and revisit different standards across the school year, focusing on interconnectedness of SEPs, CCCs, and DCIs across time, they will need a well-articulated rationale for such a change as well as instructional supports to effectively make the transition.

As design principles that guide the development of coherent instructional materials are more widely used across large-scale instructional programs, developers will need to embed supports for teachers in the materials. These supports include making visible how multiple standards are addressed in an instructional sequence, explaining the rationale for the sequencing of activities, and clarifying how each activity supports the targeted standards. Understanding the match between learning goals, standards, and learning experiences is particularly important when instructional time is short. Materials will need to provide guidance to teachers on how to modify instruction, potentially by identifying which activities, lessons, or chapters could be eliminated, with as little effect on the content storyline as possible. In other cases, materials will need to emphasize to teachers when not to eliminate content or activities, as streamlining units too much has been shown to diminish student learning (Clark & Linn, 2003; Roseman et al., 2013). In the case of large-scale programs that have influence on district curricula or pacing guides, these materials can provide recommended pacing guidelines that allow for ample time to conduct in-depth exploration around three-dimensional learning. Thus, the rationale for the extended units and the deliberate sequencing of units will need to be provided to address practitioners’ concerns about instructional time and adequate coverage of local, state, or national standards and to provide necessary insight for the intentional sequencing of units within the instructional materials.

**Recommended indicators for evaluating this criterion**

1. Do the teacher materials provide a rationale for how the science storyline is intentionally sequenced for student learning?

2. Is there guidance on how multiple standards are addressed within an instructional sequence (when applicable)?

3. If there is a pacing guide provided with the materials, does it show how the sequence of units connects to one another and builds over time?

4. Are there supports to make instructional decisions about modifications to the materials when instructional time is short (while still keeping the core activities to support the storyline)?
Given that many science teachers do not reach the “expert” level of a given scientific discipline and have not learned science in a NGSS-informed science education system, it is likely that many science teachers have compartmentalized views of science content instead of the integrated views represented in the current vision of science teaching and learning. To build towards more coherence in our curricula, therefore, many teachers need to enhance their own science knowledge and pedagogical content knowledge around three-dimensional science learning, especially with respect to science and engineering practices (NRC, 2015; Windschitl, 2002). Teacher content knowledge is one predictor of student learning. Research shows that students of content-area teachers who have a solid foundation of subject-matter knowledge tend to outperform students of teachers who have limited subject-matter knowledge (Darling-Hammond, 2000; Monk, 1994).

Teachers need continual support for integrating the knowledge they have, adding new knowledge as the fields of science and engineering evolve, and building pedagogical content knowledge (Beyer, Delgado, Davis, & Krajcik, 2009; Davis, 2004; Davis & Krajcik, 2005). When teachers readily see the interconnected and dynamic nature of science, their presentation of content to students can mirror this understanding (Brickhouse, 1990). This perspective on science promotes development of teachers’ understanding of the links between the science ideas, practices, and concepts they are teaching, and helps them enact the NGSS learning goals for students. When teachers feel confident in what they are teaching, they can more readily use that knowledge to make instructional decisions. Teacher materials have the potential to be an important resource by aiding teachers in building their own knowledge of science and science pedagogy, especially around the NGSS vision of three-dimensional learning focused on explaining phenomena or solving problems.

### Recommended indicators for evaluating this criterion

1. Do teacher materials provide instructional supports and opportunities that promote the development of content knowledge and pedagogical content knowledge in science?

2. Do the teacher materials emphasize the interconnected nature of science, including how the SEPs, CCCs, and DCIs are interconnected and supported by the learning experiences?
Category C: Instructional Materials Support Learning Experiences across the Three Dimensions.

American students report a positive attitude toward science and an interest in learning science through the early years of instruction, but this positive disposition declines in the years from middle school through high school (Osborne, Simon, & Collins, 2003; Speering & Rennie, 1996). Many argue that the changes in student interest and enjoyment of science relate to changes in teaching style (Osborne, Simon, & Collins, 2003) and perceived relevance and value for learning science (Ainley & Ainley, 2011; Speering & Rennie, 1996). Research has shown that students can remain engaged in learning science when they actively participate in the practices of science (Duschl, Schweingruber, & Shouse, 2007; Michaels, Shouse, & Schweingruber, 2007).

In addition to attitudes and interest in science, personal motivation plays a key role in student learning. Studies identify different factors that contribute to motivation (e.g., Eccles, 2005; Pintrich, Conley, & Kempler, 2003; Wigfield & Eccles, 2000). We know that students engage well when they have a sense of purpose for learning, hold an intrinsic desire to learn or a genuine interest in the topic, and see the materials as valuable and relevant (Brophy, 1999). When science content is abstract and disconnected from students’ experiences, students may perceive little value in knowing science; thus, high-achieving students may continue to excel because they value education or value a high grade, but low-achieving students will struggle because of the conceptual difficulties of the content and because they have no motivation to learn content they perceive as boring, abstract, and of little use in their lives.

Deep science learning requires meaningful problems—meaningful to students and meaningful to the discipline—because what students learn depends on both the content of the curriculum and how they encounter that content (Duschl, Schweingruber, & Shouse, 2007). In the past, observations of US classrooms have shown students engaged in activities that represent a passive and narrow view of science learning. The activities observed are often disconnected or loosely connected to concepts, and teachers tend to ask questions that require recall-level responses rather than questions that probe student thinking and experience (O’Sullivan & Weiss, 1999; Roth et al., 2006; Stigler et al., 1999). Curriculum alone does not provide engaging phenomena or meaningful problems that support deep science learning, but instructional materials can provide this context.

The criteria in this category focus on how student materials support student learning experiences and how teacher materials support teachers in using effective strategies alongside the materials themselves. Student materials ...

6S. provide multiple opportunities for students to share and negotiate their ideas, prior knowledge, and experiences.

7S. use motivating contexts to engage students in real-world phenomena and authentic design problems.

8S. are accessible to a wide range of students.

Teacher materials ...

6T. support teachers in anticipating common student ideas and include guidance to elicit and challenge student thinking.

7T. provide guidance to teachers for using effective teaching strategies that engage students in real-world phenomena and authentic design problems.

8T. provide suggestions for how to address a range of students’ skills, needs, and interests.
All students can grasp and construct insightful explanations about their world, but sometimes their ideas about phenomena are rudimentary, naive, or even wrong. Good teachers capitalize upon the rich and diverse, even if inaccurate, ideas that students bring to science class as they facilitate students’ movement to more scientifically accurate ideas. Often, however, students’ initial understanding is not engaged, and students fail to reconcile their existing knowledge with the new information they acquire through classroom instruction (Chi, 2005; Duschl, Schweingruber, & Shouse, 2007; Roth et al., 2006). Actively eliciting student ideas and revisiting them over time is a necessary step for drawing students into instruction (Duschl, Schweingruber, & Shouse, 2007; NRC, 2000).

Instructional materials can include activities that draw on students’ curiosity and questions, allowing ample opportunities for students to talk, debate, and reflect during an instructional sequence (Resnick, Michaels, & O’Connor, 2010; Roth, 2002). Without such opportunities, students’ misconceptions about science will persist, even alongside more scientific conceptions (e.g., Chi, 2005; Inagaki & Hatano, 2002). Providing opportunities for students to talk through and reconcile their thinking, particularly with one another through peer-to-peer collaboration, is important for students as they are learning science, especially because the teacher and textbook have traditionally dominated the classroom discourse. Students perceive the teacher and textbook as the “authority” of the classroom and hesitate to ask questions or voice struggles with learning (Cornelius, 2004).

Well-designed instructional materials provide opportunities for students to reveal their ideas through discourse and rich discussions, writing activities, and collaborative group activities in which students wrestle with and make sense of new information. These discussions lead to a deeper understanding of the NGSS dimensions and the nature of science. Collaborative learning encourages use of science language and authentic science discourse (BSCS, 2008; Flick, 1995; Kamen et al, 1997; Lemke, 1990), particularly for English language learners (Calderon, 1999; Stoddart et al., 2002; Stoddart et al., 2010). Instructional materials can provide ways for students to capture and revisit their thinking throughout an instructional sequence using discussion prompts, graphic organizers, and other writing strategies.

Rather than treating science as being in a “final form” (Duschl, 1990) composed of a set of facts, theories, and solved problems to be transmitted, instructional materials can treat science as a dynamic discipline. Students need to experience science as practice by asking questions and defining problems; finding ways to explore these questions empirically through research and investigation; evaluating competing, alternative models; constructing explanations and solving problems; and arguing from evidence. These kinds of learning experiences enable students to participate in science as practice and to more fully understand core concepts (Duschl, Schweingruber, & Shouse, 2007).
The importance of making science relevant and valuable to students cannot be overstated. In their work, Ainley and Ainley (2011) concluded that “when students believe that the topics they are dealing with in science have personal relevance and meaning for their lives they are more likely to experience enjoyment and interest from engaging with science content” (p. 11). As intuitive as this statement seems, traditional science instructional materials do little to emphasize relevance and connect the science ideas students are learning with explanations of phenomena in the world around them. When instructional materials do emphasize real-world contexts and applications, we see improvements in student learning (e.g., Rivet & Krajcik, 2004; Schneider, Krajcik, Marx, & Soloway, 2002) and their attitudes toward science (Siegel & Ranney, 2003). Science education has long supported experienced-based approaches to education, notably problem-based learning (e.g., Barron et al., 1998; Hmelo-Silver, 2004) and science, technology, and society (Fensham, 2009). However, weaving these authentic science experiences into instructional materials can only enhance student learning and interest if they are compelling to the students (Buxton, 2006; Lee & Songer, 2003). An authentic real-world context, for example, does little to motivate a student if the student has no interest in the topic and perceives no value for learning it (Fensham, 2009).

Instructional materials can provide a motivating context to engage students by offering contexts that are personally or socio-culturally relevant to them, fostering learning through social and collaborative interactions, using project-based or place-based approaches, and providing activities, such as an engineering design problem, that allow students to experience a challenge and the satisfaction of overcoming the challenge. These kinds of activities situated in a relevant context promote students’ engagement in the practice of science rather than experiencing science as a static set of facts or theories to be learned. Table 2.3 includes examples of phenomena or problems that may be of interest to students, although tailoring the issues and problems to students’ local community, while not always possible, could make them even more relevant.

---

**Recommended indicators for evaluating this criterion**

1. Do students have the opportunity to share their existing knowledge and experiences in relation to the topic at the beginning of an instructional unit?
   - How often are there opportunities for students share their ideas?
   - Do the materials include a way for students to capture initial ideas to revisit later?

2. To what extent do the materials emphasize revisiting student ideas when new information is presented or acquired?
   - What is the frequency with which ideas are revisited?
   - Are there opportunities to revisit these ideas in depth?

3. How well do materials embed practices of science as an integral part of learning science?

---

**When evaluating student materials, consider the extent to which ...**

**Evaluative Criterion 7S: Materials use motivating contexts to engage students in real-world phenomena and authentic design problems.**
Table 2.3. Examples of science phenomena or engineering problems.

<table>
<thead>
<tr>
<th>Science topic</th>
<th>Engaging question about a phenomenon or engineering problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metabolism</td>
<td>How does the metabolism of carbohydrates, proteins, and fats influence the performance of athletes?</td>
</tr>
<tr>
<td>Combustion</td>
<td>Why is gasoline the most common fuel we use in cars? Are there other materials we can use for fuel, like water?</td>
</tr>
<tr>
<td>Human impact on a river ecosystem</td>
<td>Can people change the way a river flows? How? And what effects would that have? How are water systems near me engineered to bring water to my community?</td>
</tr>
<tr>
<td>Invasive species</td>
<td>What effect does an increase in the zebra mussel population in the Hudson River have on the ecosystem? How are local invasive species affecting my community?</td>
</tr>
<tr>
<td>Heredity</td>
<td>How can some siblings look very similar while other siblings look very different? Are we close to being able to alter our genes for “looks” or health problems?</td>
</tr>
<tr>
<td>Design solutions</td>
<td>How can we optimize the design of windmill? What would be the best ways to harness wind energy in my community?</td>
</tr>
<tr>
<td>Science, engineering, and society</td>
<td>What effect does an increase in the zebra mussel population in the Hudson River have on the ecosystem? How are local invasive species affecting my community?</td>
</tr>
</tbody>
</table>

Recommended indicators for evaluating this criterion

1. Are the learning goals, instructional activities, and text situated in the context of a real-world phenomenon or authentic design problem that is relevant and engaging for students?
   a. Does the real-world phenomenon/design problem make connections, when appropriate, to contexts relevant to students?
   b. Do the connections attempt to make science valuable to students’ lives?
   c. Are students able to apply their new knowledge to explain other phenomena or design solutions to problems beyond the task at hand?

2. How well do the materials promote learning science as practice through relevant contexts including real-world phenomena or design problems?

When evaluating student materials, consider the extent to which ...

Evaluative Criterion 8S: Materials are accessible to a wide range of students.
Classrooms are diverse places of learning. Students bring with them rich and varied knowledge about the world influenced by their cultural and social experiences outside the classroom. There is no one-size-fits-all approach to learning science (Lynch, 2001), so instructional materials need to attend to the variety of student needs necessary for creating instructional goals, methods, and assessments that work for everyone. This will be especially important for students who are traditionally distanced from science learning (Rodriquez, 2015; Tate, Clark, Gallagher, & McLaughlin, 2008).

The NGSS clearly communicate a vision of learning for all students and calls for equitable science teaching (see especially Appendix D). A notable critique of the NGSS with respect to equity was put forth by Rodriguez (2015), who applauded the NGSS for several innovative changes to the science standards but also directed attention to the looming issues of equitable science instruction, such as the disproportional time allotted within schools to improve literacy and math performance as compared to the limited instructional time for science. Subsequently, these schools struggling in literacy and math get further behind in science. While instructional materials in science cannot fix the inequalities in instructional time, materials adopted by school districts struggling to meet language and mathematics expectations can justify more instructional time in science by focusing on literacy and math integration in science, especially around reading and comprehending content-rich texts, developing academic vocabulary, speaking and listening skills through science discussions (or science talks [Gallas, 1995]), academic writing in science, and mathematical thinking. Materials and programs that emphasize literacy and mathematics integration in science have shown gains in student learning (e.g., Guthrie et al., 2004; Lee et al., 2008; Morrow, Pressley, Smith, & Smith, 1997), especially for ELL students (Lee, 2005; Lee et al., 2005; Lee et al., 2008; Rosebury, Warren, & Conant, 1992) and learners in urban settings (Varelas & Pappas, 2006). It is especially important that such instructional programs do not lose focus on developing conceptual knowledge in science; rather, the focus of these programs is on pursuing science conceptual knowledge and engagement with science practices using literacy and mathematics skills (Pearson, Moje, & Greenleaf, 2010).

The NGSS provide a variety of strategies found to be successful for groups of students with diverse learning needs (see NGSS, Appendix D, p. 31). We know that reform-based instructional materials that use more hands-on, tactile, and manipulative activities are particularly beneficial for diverse groups of students (Geier et al., 2008; Lynch, Kuipers, Pyke, & Szesze, 2005), especially for reducing the language burden for ELL students (Lee et al., 2005). We also know that intentional and varied grouping strategies and collaborative activities support a variety of learners with different needs (Rodriquez, 2015), such as gender grouping, mixed-ability groups, and partnering ELL students with fluent bilingual speakers. More time to engage in active, participatory speech and integrating knowledge through dialogic talk is much more reflective of the home discourse experienced by some students (Bouillion & Gomez, 2001; Geier et al., 2008; Lynch, Szesze, Pyke, & Kuipers, 2007; Lynch et al., 2007) and is also a way to open the door for students’ funds of knowledge to be recognized and tapped into during classroom discourse (Moll & Gonzáles, 2004). Ample time engaging in discussion of science ideas has been shown to benefit all learners, regardless of special learning needs (Roth, 2002), and has been shown to improve students’ overall disposition toward science and view of themselves as a part of science (Warren et al., 2001). These are just a sample of research-based strategies that can be incorporated into the learning experiences put forth by instructional materials.

In addition to these strategies, instructional materials can incorporate the following:

1. Culturally relevant science and science in place (place-based examples/problems/investigations)
2. Multi-modal tasks and experiences—oral, visual, written, with various representations
3. Literacy integration with literacy goals and skills specified for students, especially ELL students
4. Rich, multicultural scenarios and vignettes, especially including non-Western and female examples
5. Accommodations for learning exceptionalities (i.e., learning disabilities, giftedness)

By drawing out students’ experientially and culturally diverse ideas and abilities, instructional materials can support more equality and access to science.

**Recommended indicators for evaluating this criterion**

1. Do the materials capitalize on diverse cultural backgrounds of students?
2. Do the materials include instructional activities and text accessible for students with learning exceptionalities (i.e., students with special needs, high-achieving students)?
3. Do the materials include multi-modal (e.g., oral, written, artistic) opportunities for students to share their thinking?
4. Do the materials include modifications and/or strategies for students with English language learning needs?
5. Do the materials help to promote gender equity and access?

**When evaluating teacher materials, consider the extent to which ...**

**Evaluative Criterion 6T: Materials support teachers in anticipating common student ideas and include guidance to elicit and challenge student thinking.**

Given the prominent role of student ideas for instructional decision-making, teacher materials should support teachers in anticipating students’ ideas and highlighting points where students’ ideas may conflict with learning goals. Dewey (1956) conceptualized this challenge as psychologizing the curriculum, arguing that master teachers were uniquely positioned to draw out student experiences and ideas as they learned new disciplinary content. Research continues to show that highly effective teachers with deep content knowledge often excel at eliciting student ideas and dealing with them as students acquire new information but that novice teachers and teachers with limited content knowledge avoid student ideas (e.g., Akerson, Flick, & Lederman, 2000). A teacher’s ability to help students come to understand SEPs, CCCs, and DCIs given students’ initial ideas is especially important to build in early career teachers and those with limited content background (Magnusson, Krajcik, & Borko, 1999). Well-designed instructional materials can support these teachers in anticipating and identifying students’ ideas that are inconsistent with current scientific understanding.

While the NGSS help direct teachers’ attention to age-appropriate SEPs, CCCs, and DCIs, instructional materials can play a vital role in helping teachers anticipate student ideas and abilities in relation to scientific learning goals. Teachers must anticipate student ideas related to both concepts and practices as well as students’ abilities to use the practices of science. Materials can do this by providing teachers with background on commonly held student ideas (from the research literature) and including activities to elicit student ideas early and often in an instructional sequence. Materials that help teachers “anticipate and interpret what learners may think about or do in response to instructional activities” are an important feature of educative curriculum materials (Davis & Krajcik, 2005, p. 5). Instructional materials that provide graphic organizers and other tools to record and track how student thinking about SEPs, CCCs, and DCIs evolve over
an instructional sequence assist students in making sense of science concepts. Importantly, given the traditional, teacher-dominated classroom discourse, instructional materials can provide discussion strategies to help teachers change their role from the leader of classroom discussion to a facilitator who draws out the student voice through true, dialogic discussions (Bahktin, 1981; Hand, Norton-Meier, Staker, & Bintz, 2006; Roth, 2002).

### Recommended indicators for evaluating this criterion

1. How well do teacher materials support teachers in anticipating commonly held student ideas across the three dimensions?
2. How well do the materials help teachers to elicit student ideas?
3. How well do the materials guide teachers in helping students negotiate and challenge their understanding?
4. Do the teacher materials point to specific moments in instruction in which student ideas should be compared or examined in relation to the learning goals?

### When evaluating teacher materials, consider the extent to which ...

**Evaluative Criterion 7T: Materials provide guidance to teachers for using effective teaching strategies that engage students in real-world phenomena and authentic design problems.**

The *Framework* and the NGSS describe the need to integrate the SEPs, CCCs, and DCIs, but teachers need support to do this well. Building teachers’ pedagogical content knowledge around three-dimensional learning is essential here. Pedagogical content knowledge (PCK) is knowledge of, reasoning behind, and planning for teaching a particular topic in a particular way for an intended purpose to enhance student outcomes (Gess-Newsome, 2015; Shulman, 1987). In many ways, PCK is implicated across all the criteria proposed for teacher materials, but this criterion focuses specifically on how teacher materials can enhance PCK around the use of phenomena and authentic design problems to promote science learning.

For the most part the science teacher workforce was educated in a system in which science content was typically taught separate from science skills and where the “scientific method” has been the dominant approach to teaching science skills or practices. Rodriguez (2015) points out, “a pre-service teacher who has mainly been exposed to traditional and transmissive pedagogy for 12–14 years and then is exposed to student-centered, hands-on, culturally relevant pedagogy for 15 weeks in a science methods course—while observing regular teachers implement transmissive approaches during student teaching—is most likely to end up mimicking what appears to be ‘the safest practices’” (p. 1,041). While teaching that supports the focus of phenomena and design problems will need to be fostered in ways that go well beyond instructional materials (e.g., teacher preparation, student teaching, coaching, and professional development), materials can play an important role. Research shows that beginning teachers rely heavily on instructional materials (Grossman & Thompson, 2004), which means that instructional materials have great potential to reinforce NGSS reform-based teaching. Practicing teachers—from early career to veteran teachers—also need support for transforming their practice around three-dimensional learning, particularly when it comes to using science ideas to explain natural phenomena and design solutions to problems.
Teacher materials can support teachers by offering a consistent set of strategies and scaffolds to support such learning across the curriculum. From previous reform efforts emphasizing inquiry science, we know that shifting practice away from the traditional topical approach is challenging when teachers are not provided explicit guidance for enacting reform practices in the classroom (Marx et al., 1998). Materials should articulate to teachers the purpose of different strategies, how the strategy supports three-dimensional learning and phenomena-focused teaching, and instructions for using the strategy and scaffolded activities across an instructional sequence. Additionally, teacher materials should include examples of student work to demonstrate how students at varying academic levels might respond to the strategy or scaffolded activity.

**Recommended indicators for evaluating this criterion**

1. Are there opportunities provided within the teacher materials to build PCK that will support a phenomena-based or an engineering design approach to learning?
2. Do the teacher materials include explicit strategies and scaffolds to implement this approach?
3. Are examples of student work included to help teachers understand how students might respond to a given strategy or scaffold?

*When evaluating teacher materials, consider the extent to which ...*

**Evaluative Criterion 8T: Materials provide suggestions on how to address a range of students’ skills, needs, and interests.**

Davis and Krajcik (2005) argue that instructional materials should support teachers’ “ability to use personal resources and the supports embedded in curriculum materials to adapt curriculum to achieve productive instructional ends” (p. 5). Brown and Edelson (2003) described this as *pedagogical design capacity*, positioning teachers as agents of change and the most knowledgeable individual for modifying materials to meet their students’ needs. The design capacity of teachers is especially important to help them adapt materials to meet diverse student needs—whether the materials are used in a rural, low-income community; a large urban district with diverse cultures, languages, and socioeconomic backgrounds; or a wealthy suburban district with ample resources. Teachers know their school and district policies and the home and community atmosphere of their students, and they have extensive experience accommodating and adapting materials to students’ needs. Good teachers also know well what is relevant to their students given the cultural and social climate of the community and the students’ background experiences (Barton, Tan, & Rivet, 2008; Rodriguez, 2015). Even the best designed instructional materials will not be able to accomplish what a good teacher can during enactment. Teachers are, therefore, uniquely positioned to ensure productive changes are made to instructional materials to benefit their students (Tate, Clark, Gallagher, & McLaughlin, 2008). Nevertheless, instructional materials can assist teachers in this process so that materials are still enacted with integrity.

From the research literature we know that the following instructional approaches and strategies in instructional materials support diverse learning needs:

- Hands-on, inquiry-oriented activities where students are “doing” science coupled with extensive opportunities to talk about how their ideas are changing. Lynch and colleagues (2005) found that inquiry
materials paired with ample talk and writing resulted in significant achievement gains for all ethnic groups as well as for students identified for free and reduced lunches and ELL students. In addition, these groups experienced positive gains in motivation and engagement.

- Students tend to engage in science learning when their experiential and cultural backgrounds are leveraged during instruction. This does not mean that instruction stays focused on students’ cultural knowledge and experiences but that it expands upon it across an instructional sequence. This is what Gay (2002) calls “cultural scaffolding—that is, using their own cultures and experiences to expand their intellectual horizons and academic achievement” (p. 109).

- Students who struggle on literacy and mathematics high-stakes assessments can improve their science achievement alongside literacy and mathematics achievement when the three are integrated thoughtfully (Lee et al., 2008; Lee, Quinn, & Valdés, 2013).

Instructional materials would do well to make discussion a central activity that is coupled with hands-on, three-dimensional learning, with talk happening before, during, and after hands-on activities (Roth et al., 2011; Warren et al., 2001). Teachers need guidance on how to implement these strategies in the context of science as well as guidance on how to support rich discussion among diverse class members as opposed to recitation that is commonly mistaken for classroom discussion (Cazden, 2001).

<table>
<thead>
<tr>
<th>Recommended indicators for evaluating this criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do the teacher materials include supports and strategies to address diverse learning needs and ensure equality and access?</td>
</tr>
<tr>
<td>2. Do the teacher materials suggest ways to capitalize on diverse cultural and social backgrounds of students?</td>
</tr>
<tr>
<td>3. Do the teacher materials include strategies for working with students with English language learning needs?</td>
</tr>
</tbody>
</table>
Category D: Instructional Materials Provide Ways to Monitor Learning across the Three Dimensions.

Assessment of NGSS-based learning goals will require substantial shifts in how assessments are written, administered, and scored. Many science assessments, such as summative chapter tests, benchmark tests, and state exams, have largely focused on conceptual knowledge and have typically assessed that knowledge through multiple-choice items. While multiple-choice items are becoming increasingly grounded in research on student learning (such as the learning progression–based items), these types of assessments are limited in their suitability for assessing SEPs and CCCs and even more limited on how well they assess integration of SEPs, CCCs, and DCIs. Computer-based assessments show promise in assessing science practices (e.g., NCES, 2012), but at present most assessments lack measurement in this area. Instructional materials need to utilize varied and creative assessment mechanisms with respect to three-dimensional learning, shifting the focus to measurement of integrated knowledge and science practices as opposed to demonstration of conceptual knowledge only. A well-designed instructional program includes a variety of assessment types that are used regularly to monitor student learning toward the learning goals (NRC, 2001).

An additional assessment innovation that should be used regularly in instructional materials is self-assessment by students. Metacognitive practices and reflection are beneficial to all students (White & Frederickson, 1998) but are notably helpful to students who perform poorly on standardized assessments (Rodriguez, 2015). As instructional materials introduce assessment of three-dimensional learning and self-assessment, teachers will need guidance on how to use assessment data to inform their practice. This is especially true for embedded formative assessments where teachers will be expected to carefully monitor student progress and be responsive in their teaching.

In order to monitor student learning in science, student materials should ...

9S. include accessible and unbiased formative and summative assessment of students’ three-dimensional learning.

10S. include multiple opportunities for self-assessment and reflection to promote sensemaking among students.

Teacher materials should provide support to help teachers interpret their students’ assessments, provide feedback to students, and use this information to plan for future instruction. Teaching materials should ...

9T. highlight formative and summative assessments and provide tools and guidance for interpreting evidence of three-dimensional learning and using assessment results to plan for instruction.

10T. provide guidance for teachers to use data from assessments to provide feedback to students and promote student self-assessment and reflection.
Pellegrino and colleagues (NRC, 2001) described a balanced assessment system as being coherent with learning goals and instruction, comprehensive in the use of varied assessment methods to capture the variety of student learning as it occurs, and continuous by showing progress over time for individual students.

Coherence between learning goals, instruction, and assessment is always the aim of good instructional programs; therefore, assessments must reflect what students are doing and learning during instructional sequences. With the focus on integrated knowledge and practice called for by the NGSS, summative and formative classroom assessments will need to be revised and/or created to assess the integrated goals that drive instruction. While some SEPs, CCCs, and DCIs could be assessed separately, for the most part assessments should seek to assess at least two or three dimensions together. This relates to the NGSS Performance Expectations. While the Performance Expectations are not the only example of three-dimensional learning, they are exemplars that can guide assessment design. Good assessment of these tasks will utilize diverse and varied assessment methods. For example, most traditional assessments are inadequate for assessing a student’s ability to construct an argument from evidence, to develop a model, or to plan and conduct an investigation (NRC, 2012). Effective assessments of three-dimensional learning will likely employ performance assessments to adequately capture student learning (Darling-Hammond, 2014; NRC, 2014).

Instructional materials should employ a comprehensive array of assessment tasks that call for different ways to exhibit learning, with special attention to eliminating “barriers of language, gender-biased examples, and other forms of representation that preclude some students’ useful participation” (NRC, 2012, p. 261). Using more than a one-time snapshot of learning provides a more valid and fair measure of student learning, especially for those students who do not perform well on traditional assessment measures. One highly effective form of assessment becoming more widespread in classrooms is embedded formative assessment, which has been shown to improve student learning when students are provided clear, specific, and timely feedback from the assessment (Black & Wiliam, 1998; Sadler, 1998; Wiliam & Leahy, 2007). A variety of formative assessments, combined with summative assessments, can allow for multiple and varied opportunities for students to demonstrate their understanding of the SEPS, CCCs, and DCIs.

Instructional materials should also focus on continuous assessment across time for individual students, especially as it relates to students acquiring increasingly more sophisticated knowledge and practice over an instructional sequence(s). For example, one can imagine that using an assessment of “arguments from evidence” periodically throughout the school year would reveal increasingly sophisticated arguments as students learn argumentation across several instructional units and become more adept at engaging with this practice. Assessment of student learning of CCCs across instructional units is one area in which continuous assessment will be of particular use. Instructional materials that include suggestions for coherent, comprehensive, and continuous formative and summative assessments will support teachers in using a balanced assessment approach in their instruction.
A metacognitive ("thinking about thinking") approach to instruction can help students take control of their own learning by engaging them in understanding learning goals and monitoring their progress in achieving them (NRC, 2000). One way to support metacognition and reflection among students is to use self-assessment tasks and activities (NRC, 2014). Metacognitive and reflective assessments have been shown to benefit all students but are especially beneficial for low-achieving students (Black & Wiliam, 1998; Rodriguez, 2015; White & Frederickson, 1998, 2000). Through metacognition, students reflect on their role in inquiry and on the monitoring and critiquing of one’s own claims as well as claims made by other students. Black and Wiliam argue that, “if formative assessment is to be productive, pupils should be trained in self-assessment so that they can understand the main purpose of their learning and therefore grasp what they need to do to achieve” (p. 143). Thus, self-assessment can empower students to take charge of their own learning.
Student materials should include student-friendly tools for monitoring their ideas across time, such as “thought trackers” (Windschitl, Thompson, Braaten, & Stroupe, 2012, p. 893) or self-monitoring tools developed by Web-based Inquiry Science Environment (2015). Metacognitive tools allow students to be aware of what they know or do not know, understand where they are in a work process, and pinpoint their confusions.

Some examples of self-assessment include

- instructions to students to look back at what they thought before the activity and ask questions such as the following:
  - How has my thinking changed?
  - What was responsible for the changes in my thinking?
  - How did [a particular tool or strategy] influence my thinking?

- structures or scaffolds to help students develop metacognitive abilities. This might include a sentence starter such as the following:
  - My model provides a causal account of the phenomena because it shows how ...

- structures that highlight the metacognition of others. For instance, consider prompts such as the following:
  - What is the difference in thinking between Student A’s and Student B’s explanations?

### Recommended indicators for evaluating this criterion

1. Do the materials include tools for students to track their understanding at the beginning, middle, and end of an instructional unit?
   a. Is there a clear pre-assessment task at the start of the unit where students record what they know?
   b. Do the materials include opportunities for students to revisit their ideas at one or more points during instruction, with specific prompts for reflecting on their initial ideas and identifying points of confusion?
   c. Do the materials include a reflection at the end of the instructional sequence for students to articulate how and why their ideas changed over time?

### When evaluating teacher materials, consider the extent to which …

**Evaluative Criterion 9T:** Materials highlight formative and summative assessments and provide tools and guidance for interpreting evidence of three-dimensional learning and using assessment results to plan for instruction.
Assessments for learning can be used by both students and teachers to check progress toward learning goals, but to promote the productive use of assessment in the classroom, teacher materials need to highlight the embedded assessments and the learning goals they target and provide tools and guidance for how to interpret responses to those embedded assessments for planning for future instruction.

Within any given instructional sequence, there may be many different types and grain sizes of assessments. For example, questioning strategies embedded in an activity could help teachers promote student discourse to gauge their initial understanding. Other possible assessments include questions for written or oral responses, quizzes, or a unit exam (NRC, 2014). It is important for materials to highlight the assessments of student learning and their purpose (e.g., formative, summative), paying particular attention to the assessments that target three-dimensional learning. Materials should also call out the alignment to the learning goals so that teachers can track their students’ progress toward the goals that are part of an instructional sequence (NRC, 2014).

Each assessment item should include tools for interpreting the results. These tools might simply be examples of ideal responses and possible student misunderstandings, but they might also include rubrics with scoring instructions to assist teachers in determining the results of the assessment. Further, these tools should help teachers not only score or grade student work but also interpret what student answers may mean in terms of understandings or misunderstandings. Guidance for interpretation should “clearly define the forms of evidence associated with beginning, intermediate, and sophisticated levels of knowledge and practice expected for a particular instructional sequence” (NRC, 2014, p. 91). Information about where students should be at the start and end of an instructional sequence can help teachers gauge student progress. Assessment tools could also provide productive intermediary prompts (NRC, 2014). This guidance should identify key decision points within an instructional sequence, point out ideas students must understand to be successful in the next lesson, or provide suggestions of additional activities or learning strategies to make student thinking visible and help them negotiate their understanding. With appropriate support from materials, teachers will be able to accurately interpret student assessments in order to give feedback to students and to make instructional decisions responsive to student learning.

**Recommended indicators for evaluating this criterion**

1. Are tools provided for scoring assessment items (e.g., sample student responses, rubrics, scoring guidelines, or open-ended feedback)?

2. Is guidance provided for interpreting the assessments (i.e., determining what high or low scores mean for student learning)?
   a. Does this interpretation identify the level of understanding students have for the relevant SEPs, CCCs, and DCIs?

3. Are explicit strategies given for using the results of embedded assessments to plan for future instruction?
Pellegrino and colleagues (NRC, 2001) state, “The power of classroom assessment resides in its close connections to instruction and teachers’ knowledge of their students’ instructional histories” (p. 220). The power of classroom assessment is also in the ability to provide feedback to students about their learning. Feedback allows teachers and students to better understand what they know and what they are unsure of or confused about, which can be built on to lead to future learning.

However, Black and colleagues (2004) found that “assessment feedback often has a negative impact, particularly on low-achieving students, who are led to believe that they lack ‘ability’ and so are not able to learn” (p. 8). This research found that the quality of comments given as feedback to students on their assessments had a statistically significant effect on student learning compared to numerical scores. In fact, numerical scores had a negative impact on students and actually caused students to spend less time reading through comments and thinking about ways to improve their work. Focusing more on commenting on student work (as opposed to providing a numerical score) enabled teachers to glean more about student understandings and misunderstandings. They also reconsidered what was being assessed and how important it was for student learning. These findings are also echoed in Dweck’s (2008) work on growth mindsets. She found that students who had more of a growth mindset and desire to learn, rather than a strong focus on grades, were able to improve their achievement in math and science, whereas students with a fixed mindset toward their intelligence in math and science had decreasing achievement. Teachers play a key role in “shaping students’ mindsets” and can encourage a growth mindset by providing feedback to students in ways that sustain growth rather than discourage it (p. 2, 7–8).

Instructional materials can provide suggestions to teachers on how to provide feedback to students during the learning process, especially on assessments. To do this well, materials can use feedback strategies and allow for the time needed for teachers and peers to review work and for students to reflect on the comments they receive, revise their thinking, and even respond to those who provided feedback. Black and colleagues (2004) emphasized the role of written tasks and time by articulating three approaches to improving assessment feedback for students:

- Written tasks, alongside oral questioning, should encourage students to develop and show understanding of the key features of what they have learned.
- Comments should identify what has been done well and what still needs improvement and give guidance on how to make that improvement.
- Opportunities for students to respond to comments should be planned as part of the overall learning process. (p. 14)

**Recommended indicators for evaluating this criterion**

1. Are teachers given guidance on how to provide feedback to students about their learning?
2. Does guidance regarding feedback include multiple forms of feedback beyond numerical scores?
3. Do materials suggest instructional activities in which the teacher and students reflect on feedback and/or respond to the feedback they received?
Concluding Remarks

The criteria presented in this chapter represent the ambitious vision of science education put forth by the NRC Framework and the NGSS along with additional research from the science education community about effective instructional materials. In developing these criteria, it became apparent that both curriculum developers and evaluators will need to consider two major issues in light of the NGSS.

The first is the emphasis on three-dimensional learning within the NGSS along with the additional connections to the nature of science; engineering, technology, and the applications of science; and Common Core ELA and math. Incorporating this dynamic and multifaceted view of science into instructional materials will be a considerable challenge. The criteria put forth in this chapter are meant to support the SEPs, CCCs, and DCIs and the integration of these three dimensions throughout an instructional sequence. Not all activities within an instructional sequence will include all three dimensions, but rather, designers of instructional materials should carefully select and integrate the dimensions as appropriate to support deep learning of science in meaningful ways (see NRC, 2012, pp. 217–219 for more information on integrating the three dimensions). Three-dimensional learning is threaded throughout all the categories of criteria as a result of this important emphasis of the NGSS. The evaluative criteria also include significant connections to the nature of science; applications of science, engineering, and technology; and English language arts and mathematics standards as called for by Common Core State Standards. These connections are also threaded throughout the categories of criteria, providing students rich opportunities to experience science in more authentic ways by integrating relevant aspects of other subject areas. Finally, these criteria are written with the vision that science learning should be accessible and engaging to all students, valuing their ideas and experiences and providing ways for them to reflect on their learning of science. Thus, it is the intent that all categories of criteria support equitable learning experiences for all students.

The second issue is that of the design decisions that occur when developing and revising materials. The evaluative criteria described in this chapter represent a close connection to the vision put forth by the NGSS; however, it is important to note that given the realities of both designing and evaluating instructional materials, no single program is expected to perform equally well on all the evaluative criteria. There are trade-offs in the design of instructional materials, and some criteria are emphasized over others within an entire program or specific parts of a program. We argue that the goal of designing instructional materials that embody NGSS is not necessarily in checking off all the boxes and meeting each criterion equally well but rather developing a better understanding of the alternatives in design decisions when some criteria are (or are not) met. In this spirit, instructional materials designers should be able to justify the necessary trade-offs given the practicalities of ensuring materials are usable and feasible for classrooms and also meet the intended purposes. Further, the expectation for meeting these criteria within a program also depends on the scope of the program; a full-year program is expected to include all these criteria to some degree, whereas a shorter curriculum module (e.g., a two-week supplementary module) might not be as comprehensive. Finally, by their nature, some criteria presented here are best evaluated at the level of a lesson or instructional sequence, whereas other criteria are best evaluated using a larger sample from the materials (i.e., across multiple units of instruction). These and other considerations for the process of evaluating instructional materials are discussed in chapter 3.
To serve the goals of the NRC *Monitoring Progress* report, a system for evaluating instructional materials should be valid and reliable. In this chapter, we consider the challenge of systematically applying the Evaluative Criteria in chapter 2 to reach valid and reliable conclusions about the quality of instructional materials and their alignment with the vision of the NRC *Framework* and NGSS. Specifically, we offer nine guidelines for designing a system to evaluate instructional materials. The guidelines are summarized in table 3.1.

Table 3.1. Guidelines for assessing the quality of instructional materials.

<table>
<thead>
<tr>
<th>The evaluation system</th>
<th>The evaluation system should include Processes that</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. includes both tools and processes.</td>
<td>7. identify appropriate units of analysis.</td>
</tr>
<tr>
<td>2. includes a guide for evaluators.</td>
<td>8. involve dialogue and consensus-building among a team of evaluators.</td>
</tr>
<tr>
<td>3. specifies a summary report that justifies the evaluation results and offers suggestions for modifying instructional materials to enhance their quality.</td>
<td>9. assure consistency across evaluators.</td>
</tr>
<tr>
<td>4. specify what to look for as evidence for each Evaluative Criterion.</td>
<td></td>
</tr>
<tr>
<td>5. have clearly defined scoring guidelines for capturing evidence from materials.</td>
<td></td>
</tr>
<tr>
<td>6. include forms for documenting specific evidence of the Evaluative Criteria and suggestions for improvement.</td>
<td></td>
</tr>
</tbody>
</table>

Examples of how the guidelines are operationalized in existing evaluation systems are used throughout this chapter. The examples are drawn from the evaluation systems listed in table 3.2.
Note: This table does not include the Heuristics for Educative Curriculum Materials measure because the heuristics were not initially written for scoring materials but rather as recommendations of features to include in materials to support teacher knowledge and implementation. They were later operationalized for use in scoring eight biology programs (Beyer, Delgado, Davis, & Krajcik, 2006; 2009).

* AIM is not currently available to the public. It is used in professional development programs.

Table 3.2. Existing evaluation systems referenced in chapter 3.

<table>
<thead>
<tr>
<th>System</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAAS Project 2061 Instructional Materials Analysis Procedure (AAAS Project 2061 measure)</td>
<td>AAAS, 2005; Roseman, Kesidou, &amp; Stern, 1997</td>
</tr>
<tr>
<td>Analyzing Instructional Materials (AIM Tools and Processes)*</td>
<td>BSCS, 2012; Powell, Short, &amp; Landes, 2002</td>
</tr>
<tr>
<td>Educators Evaluating the Quality of Instructional Products: NGSS (EQuIP: NGSS)</td>
<td>Achieve, 2016</td>
</tr>
<tr>
<td>Primary Evaluation of Essential Criteria for Alignment (PEEC-Alignment)</td>
<td>Achieve, 2015b</td>
</tr>
</tbody>
</table>

Guideline 1: The evaluation system includes both tools and processes.

Evaluation of instructional materials requires both tools and processes. Evaluation tools provide focused tasks to help evaluators analyze the extent to which the Evaluative Criteria (see chapter 2) are present in the materials. These tools enable evaluators to make specific and consistent assessments of and recommendations for instructional materials regarding the extent to which they meet NGSS goals. Specific processes are also needed for consistent evaluation of instructional materials. These processes should specify how evaluation teams will work, that is, how they will apply the tools appropriately and consistently, build consensus among team members, and establish interrater reliability. Tools and processes must be combined in a way that results in meaningful distinctions among different sets of materials.

Examples of Evaluation Systems with Tools and Process

Table 3.3 shows examples of two evaluation systems that include both tools and processes. AIM directs evaluators to construct a graphic that depicts the flow of concepts in a unit of instruction, as well as the process for doing this. EQuIP: NGSS specifies that evaluators use a rubric in a process for identifying evidence that the instructional materials support three-dimensional learning.
An evaluator’s guide conveys the overall purpose of the evaluation system as well as the specific purposes of each tool and process and how they contribute to the overall goals of the assessment. Key components of an evaluator’s guide include the following:

- The goal or purpose of the evaluation system
- A description of the relevant tools and processes
- Examples of scoring decisions
Practice exercises for applying the system

The purpose of each tool and process should be related to relevant Evaluative Criteria. An evaluator’s guide specifies the evidence to collect for each criterion and where in the instructional materials that evidence is likely to be found (for example, in the student materials or the teacher materials). Each tool and process should include a rating scale to indicate the strength of the particular criterion, or set of criteria, in the materials. Examples from instructional materials that illustrate ratings and the rationale for those ratings are especially helpful. Ideally, the evaluator’s guide would include a range of examples, from those showing strong evidence of a particular criterion to those that have weaker evidence for it.

A different set of examples that have been rated by experts should also be included, along with the rationale for the experts’ rating. These examples allow new evaluators to practice using the tools and process within the evaluation system, comparing their ratings with those of the expert evaluators and reading and reflecting on the rationales provided by the experts.

Example of an Evaluator’s Guide

The Professional Learning Facilitator’s Guide is an online evaluator’s guide that accompanies EQuIP: NGSS (Achieve, 2015a). The guide includes a series of ten modules to support educators and education leaders in using the system. The guide also includes downloadable PowerPoint slides, key points for each slide, and downloadable handouts for each module, as needed.

For example, one slide in the third module is shown in figure 3.1. The guide provides the following key points for this slide:

- Think of the three components of three-dimensional learning as three intertwining strands of a rope. While the rope can be separated into its three different strands, the strength of the rope is determined by the strands working together; separating the strands weakens the rope so that it is no longer effective for our intended use.

- Likewise, while in the past we may have separated out the knowledge and skills students need in the study of science, in actuality, knowing and doing cannot be separated if our goal is the kind of usable, conceptual understanding students need to think, act, and learn like scientists.

- Three-dimensional learning—practices, core ideas, and crosscutting concepts working together—is therefore nonnegotiable for NGSS lessons and units.

Figure 3.1. Slide 4 from Module 3 of the EQuIP Professional Learning Facilitator’s Guide (Achieve, 2015a).
A summary report provides a quick overview of a program that helps science education leaders determine how well particular instructional materials align with the NGSS goals and values. Thus, an evaluation system should specify a summary report that describes and provides evidence of the quality of the materials and also recommends modifications for enhancing their quality. This report should include statements that relate the strength of the Evaluative Criteria within the instructional materials, allowing for easy comparison among different sets of instructional materials. In addition, the summary should include evidence that clarifies and illustrates the material's strengths and limitations as well as suggestions for relevant modifications. Because the ultimate goal of the NGSS is to improve science learning, suggestions for modification will support writers, publishers, and curriculum developers in making revisions that better exemplify the NGSS. Furthermore, such a summary can provide insights for teachers and professional development providers about how to adjust instruction in the interim to accommodate the limitations identified in previous or newly adopted materials.

Example of Process for Completing a Summary Report

The final step in the PEEC-Alignment evaluation system specifies the completion of a summary for each NGSS innovation. Figure 3.2 shows the process for creating a summary for the degree to which lessons in the instructional materials integrate the three NGSS dimensions.

SUMMARY AND RECOMMENDATIONS

Use the evidence from the program under review for the summary relative to this innovation. Answer the following questions.

1) To what degree does the program meet the criteria for integrating three dimensions?
   - Materials incorporate the innovation.
   - Materials partially incorporate the innovation.
   - Materials do not incorporate the innovation.

2) Do the materials meet an adequate level of acceptance?  YES  NO

3) If the materials meet an adequate level, describe specific changes that would improve the program further.

Figure 3.2. Example of a summary report from the PEEC-Alignment measure (Achieve, 2015b, p. 40).
Guideline 4: The evaluation system should be supported by tools that specify what to look for as evidence for each Evaluative Criterion.

Evaluators of instructional materials are best supported when tools provide guidance regarding what to look for as indicators of each Evaluative Criterion. Tools might specify, for example, what counts as evidence of three-dimensional learning or of opportunities for students to assess their own learning.

Examples of Tools that Specify Evidence for Criteria

The AAAS Project 2061 system includes a set of indicators for each criterion evaluated. Using these indicators, evaluators determine the extent to which each criterion is present in the materials. For example, one criterion specified is “Assisting teachers in identifying their students’ ideas.” Indicators are shown in figure 3.3.

**Assisting Teachers in Identifying Their Students’ Ideas**

*Indicators of meeting the criterion*

1. Material includes *specific* questions or tasks that could be used by teachers to identify students’ ideas.

2. The questions/tasks are likely to be *comprehensible* to students who have not studied the topic and are not familiar with the scientific vocabulary.

3. The questions/tasks are identified as serving the purpose of identifying students’ ideas.

4. The material includes questions/tasks that ask students to *make predictions* and/or *give explanations* of phenomena (rather than focus primarily on identifying students’ meanings for terms).

5. The material suggests how teachers can *probe* beneath students’ initial responses to questions or *interpret* student responses (e.g., by providing annotated samples of student work).

**Figure 3.3.** Indicators for “Criteria Used in Evaluating the Textbooks’ Quality of Instructional Support” from the AAAS Project 2061 Evaluation (2005).

A second example is depicted in table 3.4, which is an excerpt from the *EQuIP*: NGSS system that describes criteria for evaluating how well a lesson or unit aligns with the NGSS.

**Table 3.4.** Selected criteria from *EQuIP*: NGSS for evaluating alignment to the NGSS (Achieve, 2016, p. 2).
Both examples provide guidance to evaluators on specific indicators of how well materials meet the criterion being evaluated.

Guideline 5: The evaluation system should be supported by tools that have clearly defined scoring guidelines for capturing evidence from materials.

Tools can be advantageous for capturing evidence of how well instructional materials achieve each Evaluative Criterion (or set of Criteria). The most common tools for capturing evidence in existing evaluation systems include the following:

- Rubrics – sets of consistent criteria for evaluating a particular aspect or component of instructional materials
- Graphic Organizers – representations that depict and organize evidence gathered from instructional materials based on Evaluative Criteria.

The evidence captured by tools like these can be used to support decisions about the quality of instructional materials and the extent to which they exemplify the NGSS. This evidence can also be used to add examples to a summary report and to provide formative feedback for developers.

Examples of Tools for Capturing Evidence

AIM tools include both rubrics and a graphic organizer. Two examples are given in table 3.5, which shows a portion of a rubric that is used to evaluate the quality and use of assessments in instructional materials.

Table 3.5. Excerpt from a rubric in the AIM evaluation system (BSCS, 2012).
### ASSESSMENT RUBRIC

<table>
<thead>
<tr>
<th>QUALITY</th>
<th>(5)</th>
<th>(3)</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-quality assessments</td>
<td>The assessments have most of the features of high-quality assessments.</td>
<td>The assessments have some of the features of high-quality assessments.</td>
<td>The assessments have few of the features of high-quality assessments.</td>
</tr>
<tr>
<td>• measure what students know and are able to do</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• align with learning goals and the mode of instruction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• stress application of what students know and are able to do in new or different situations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• provide students the opportunity to assess their own learning</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### MULTIPLE MEASURES

Examples of assessments include the following:

- Performance tasks
- Objective assessments
- Constructed-response questions
- Project-based tasks
- Portfolios

A wide variety of assessment measures and corresponding scoring guidelines (e.g., rubrics, answer keys) are provided.

Some variety of assessment measures is provided.

Assessments are limited to a few different types.

---

The Conceptual Flow Graphic (CFG) from AIM is a graphic organizer used to evaluate the coherence of instructional materials. Figure 3.4 shows an example of a CFG. Evaluation teams construct the CFG by identifying the overarching concept for a chapter or unit and the major concept in each of its sections. Next, they place the concepts for each section in sequence, with the overarching chapter or unit concept in the center. Finally, the teams indicate the strength of the link between each concept and the overarching concept, as well as between sequential concepts, with thick, thin, or dashed arrows.
Examples of Tools with Clearly Defined Scoring Guidelines

Well-defined scoring guidelines are an important component of tools. For example, a tool may use a rating scale, a “scale whose points are defined by predetermined criteria (verbal, numeric, or symbolic descriptors) and with which judgments concerning the strength of a particular trait, along a continuum, are indicated” (Joint Committee on the Standards for Educational and Psychological Testing, 2014, p. 22). A rating scale supports gathering and/or quantifying subjective judgments about materials, enhancing the likelihood of similar ratings of the same materials by different evaluators. Rating scales assist evaluators in specifying the extent to which the Evaluative Criteria are represented in instructional materials.

Table 3.6 contains two examples of scoring guidelines. The AAAS Project 2061 process specifies indicators that must be met for the criterion “assisting teachers in identifying students’ ideas” in order to classify instructional materials as excellent, satisfactory, or poor (n.p.). The PEEC-Alignment process includes specific criteria for classifying the evidence related to the integration of the three dimensions of the NGSS as excellent, adequate, inadequate, or no evidence.
Table 3.6. Examples of rating scales in two existing evaluation processes.

<table>
<thead>
<tr>
<th>Evaluation process</th>
<th>Sample scoring guide</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAAS Project 2061</td>
<td><strong>Rating Scheme for Criterion “Assisting teachers in identifying students’ ideas”</strong></td>
</tr>
</tbody>
</table>
| (AAAS, 2005)              | *Excellent:* The material provides a sufficient number and variety of questions/tasks that meet indicators 1 and 2 and meet indicators 3–5.  
*Satisfactory:* The material provides some questions/tasks that meet indicators 1–4.  
*Poor:* The material provides some questions/tasks that meet indicator 1 or indicators 1 & 2. |
| PEEC-Alignment            | **Scoring Guide for Integrating Three Dimensions**                                   |
| (Achieve, 2015b, p. 25)   | *No Evidence:* This is self-evident. You cannot find any evidence for the NGSS innovation.  
*Inadequate Evidence:* You can identify one or two instances of the innovation, but they do not constitute adequate time or opportunity for students to learn the content or develop the ability.  
*Adequate Evidence:* You can identify three or four instances of the innovation, and they constitute adequate time and opportunity for average students to learn the content and develop the abilities.  
*Excellent Evidence:* You can identify five or more instances of the innovation, and they constitute adequate time and opportunity for most students to learn the content and develop the abilities. |

Guideline 6: The evaluation system should be supported by tools for documenting specific evidence of the Evaluative Criteria and listing suggestions for improvement.

Tools for recording evidence provide documentation of the extent to which the Evaluative Criteria are present in instructional materials. Once completed, these tools also provide examples that illustrate not only the criteria that are well supported by the materials but also those that could be strengthened. The evidence recorded in these tools also promotes discussion among teams of reviewers as they build consensus on the evaluation.

Examples of Tools for Recording Evidence and Suggestions

*EQuIP: NGSS* includes a form for recording evidence of desired criteria and suggestions for improving the lesson or unit (see table 3.7). The process further specifies that all feedback and suggestions be specific and based on individual criteria within each of the three categories in the measure.
Table 3.7. A portion of the form for recording evidence and suggestions from the EQuiP: NGSS measure (Achieve, 2016, p. 6).

<table>
<thead>
<tr>
<th>Lesson and unit criteria</th>
<th>Specific evidence from materials (what happened?/where did it happen?) and reviewer’s reasoning (how/why is this evidence?)</th>
<th>Evidence of quality?</th>
<th>Suggestions for improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Three dimensions: Builds understanding of multiple grade-appropriate elements of the science and engineering practices (SEPs), disciplinary core ideas (DCIs), and crosscutting concepts (CCCs) that are deliberately selected to aid student sensemaking of phenomena and/or designing of solutions.</td>
<td>Document evidence and reasoning, and evaluate whether or not there is sufficient evidence of quality for each dimension separately.</td>
<td>☐ None ☐ Inadequate ☐ Adequate ☐ Extensive (All three dimensions must be rated at least “adequate” to mark “adequate” overall.)</td>
<td></td>
</tr>
<tr>
<td>i. Provides opportunities to develop and use specific elements of the SEP(s).</td>
<td>☐ None ☐ Inadequate ☐ Adequate ☐ Extensive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. Provides opportunities to develop and use specific elements of the DCI(s).</td>
<td>☐ None ☐ Inadequate ☐ Adequate ☐ Extensive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iii. Provides opportunities to develop and use specific elements of the CCC(s). Evidence needs to be at the element level of the dimensions (see rubric introduction for a description of what is meant by “element”).</td>
<td>☐ None ☐ Inadequate ☐ Adequate ☐ Extensive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PEEC-Alignment has a similar form for recording evidence and listing suggestions, though it is organized differently (figure 3.5).

---

**CLAIM**

**TITLE OF PROGRAM:** ____________________________ provides adequate and appropriate opportunities for students to meet the Performance Expectations in the NGSS.

**EVIDENCE-BASED RESPONSE**

**RECOMMENDATIONS**

---

Figure 3.5. The PEEC-Alignment form for recording evidence and suggestions for instructional materials (Achieve, 2015b, p. 41).
A critical step in analyzing instructional materials is breaking them down into appropriate units of analysis. This will not be simple because the evaluation of different criteria requires different units of analysis. For example, determining the degree to which learning goals are made explicit to students in instructional materials could be evaluated in a selection of activities from different points in a program; however, determining the degree to which learning experiences are sequenced in a coherent storyline requires evaluating at least a chapter or sequential set of lessons. Similarly, instructional materials designed for different grade bands may require different units of analysis. Secondary school materials tend to be organized as chapters within units, while elementary school materials often take the form of lessons within a unit or module. An evaluation system should provide a process for sampling the units of analysis to be reviewed. For instance, the process might specify selecting chapters from the beginning, middle, and end of the program. In addition, the process may indicate that specific units or learning goals be the focus of the evaluation.

Examples of Processes that Specify Units of Analysis

Two existing evaluation processes that specify the section(s) of materials to be evaluated for particular qualities are the AAAS Project 2061 evaluation and the draft PEEC-Alignment evaluation (see table 3.8).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Attribute identified in measure</th>
<th>Unit of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAAS Project 2061 (AAAS, 2005, Footnote 2)</td>
<td>Makes purpose explicit to students</td>
<td>Footnotes for Criteria Used in Evaluating the Programs’ Quality of Instructional Support <a href="http://www.project2061.org/publications/textbook/hsbio/report/feet.htm#1">http://www.project2061.org/publications/textbook/hsbio/report/feet.htm#1</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If a material frames sections within a unit (rather than the whole unit), then the reviewer should evaluate and rate the purpose for each section separately and take the average.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If a material attempts to frame both the unit overall and large sections within the unit, the reviewer should evaluate and rate each purpose separately. The overall rating of the criterion will be the average of the rating of the purpose for the unit overall and the average rating of the purposes provided for the individual sections.</td>
</tr>
<tr>
<td>PEEC-Alignment (Achieve, 2015b, p. 25)</td>
<td>Integrating three dimensions</td>
<td>Sample three sequences of instruction consisting of four to five activities per sequence.</td>
</tr>
</tbody>
</table>
The *Monitoring Progress* report requires a valid and reliable evaluation system for assessing how well instructional materials embody the NGSS. Thus, it is important that the evaluation of instructional materials does not rest with just one individual. Evaluators should work in teams to identify and defend to each other evidence for the Evaluative Criteria in the materials. The discussion and negotiation required to build consensus around what counts as evidence results in a solid, defensible assessment of the extent to which those criteria are present in the materials. A system that requires collaboration among a team of evaluators is a rigorous method for determining the presence of the Evaluative Criteria in instructional materials.

**Examples of Processes that Specify Use by Collaborative Teams**

Most evaluation processes reviewed in this report recommend that a team of at least two individuals evaluate the materials. Although use by individuals is possible for the *EQuIP: NGSS* measure, it is not preferred. Table 3.9 reports the recommendations for the processes reviewed here.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Statements regarding the process for evaluating materials</th>
</tr>
</thead>
</table>
| AAAS Project 2061 (AAAS, 2005, About the Evaluation) | ● *The instruction in each textbook was examined meticulously by a total of eight instructional analysts who worked as four independent two-member teams.*  
● *After the teams completed their analysis of the curriculum materials independently, the teams then met to discuss and reconcile their findings, while getting feedback from Project 2061 staff.*  
● *Thus, the ratings in the evaluation reports represent a consensus of the review teams and Project 2061 staff.*                                                                 |
| AIM Tools and Process (K-12 Alliance 2007, p. 4) | ● *It is not something that can be conducted in isolation, but requires the collaboration of teachers and administrators.*  
● *While other processes tend to focus on using a series of standards or benchmark-alignment checklists that may be done by separate committees or individuals working in isolation, the AIM process is unique because it engages all teachers—who will be the ones ultimately using the instructional materials in their classrooms.*  |
| EQuIP: NGSS (Achieve, 2016, p.4) | ● *While it is possible for the rubric to be applied by an individual, the quality review process works best with a team of reviewers, as a collaborative process, with the individuals recording their thoughts and then discussing with other team members before finalizing their feedback and suggestions for improvement.* |

**Table 3.9.** Recommendations regarding teams in evaluation processes.
Guideline 9: The evaluation system should include a process that assures consistency across evaluators.

Consistency across evaluators—interrater reliability—is not an issue within evaluation teams because, as noted in Guideline 8, the process should require teams to work to consensus on scores. It is expected, however, that to review the diverse set of widely used science programs in the United States, it will be necessary for multiple teams to evaluate instructional materials. To ensure that different teams follow the process and apply the tools consistently, a process for establishing and maintaining interrater reliability across teams is needed. Developers should consider including excerpts from sample instructional materials for each grade band and science discipline that can be evaluated by each evaluation team. The results of these practice evaluations could then be compared across teams as well as with expert evaluations. This process could determine how consistently an evaluation system’s tools and processes have been applied.

Example of Process for Establishing Interrater Reliability

In figure 3.6, the developers of the AAAS Project 2061 evaluation system describe how they established interrater reliability in a study that used this evaluation system. Their process included the use of a “practice textbook” similar to the suggestion above.

Reliability Study: As part of the high school biology textbook evaluation, Project 2061 carried out a reliability study, in which we examined the consistency of reviewers’ judgments about the textbooks’ coherence for the topic Matter and Energy Transformations. Coherence was defined by the extent to which a textbook aligned with the complete set of key ideas for the topic and whether it made explicit connections among the key ideas. Reliability in evaluating textbook coherence is an important first step for designing coherent curriculum materials. If review teams can agree with and adhere to a way of operationalizing curriculum coherence, then it becomes possible to design materials by those same guidelines.

Using a strand map adapted from Atlas of Science Literacy . . . we specified the connections among key ideas that would be used to define coherence and organized them into three categories: Connections among key ideas, Connections between key ideas and their prerequisites, and Connections between key ideas and related ideas. Reviewers used this map to keep track of their judgments and to facilitate discussion within and between review teams. The study findings indicate that the coherence of the high school biology textbooks can be consistently judged using these maps. Several factors probably contributed to the reliability, including (a) providing written clarification of the meaning of the key ideas and what “alignment” and “connections” mean for each one, (b) using knowledgeable and experienced review teams, and (c) providing training that included applying the written clarification to a practice textbook.

http://www.project2061.org/publications/textbook/hsbio/report/about.htm

Figure 3.6. Process for establishing interrater reliability for AAAS Project 2061 evaluation system (AAAS, 2005).
Concluding Remarks

This chapter includes many examples of tools and processes from the evaluation systems studied for this report (table 3.2). While these examples are intended to be illustrative of the features of evaluation systems that evaluate instructional materials, developers should not be limited by the examples provided here. For example, there may be multimedia tools that were not described in this chapter that are more effective in evaluating particular aspects of materials.

It is neither likely nor expected that one evaluation system can be used successfully to evaluate all science instructional materials. There will be differences in systems designed for materials intended for different grade bands and/or for different science disciplines. Nevertheless, any process developed for the purpose described in the Monitoring Progress report should follow the nine guidelines described in this chapter. They provide guidance for developing rigorous and consistent evaluation systems across different instructional programs and by different reviewers. These guidelines show that the tools and processes must work together to gather evidence for the quality of the materials, provide opportunities for evaluators to discuss and build consensus among reviews, and provide a neutral, equal platform to compare the strengths and limitations of the various instructional materials used in science education.
References


Brown, M., & Edelson, D. (2003). Teaching as design: Can we better understand the ways in which teachers use materials so we can better design materials to support their changes in practice [Design brief]. Evanston, IL: Center for Learning Technologies in Urban Schools.


