CHAPTER 1

Assessment for Learning and Teaching

Deciding to become a teacher is a major career decision. With much excitement and anticipation after your acceptance into a teacher education program, your journey toward becoming a science teacher starts in university classrooms. You are taking various courses, including science teaching methods courses. You may not give much attention to science assessment because your teacher education program may not offer a stand-alone science assessment course. From your grade school to university, you have seen your teachers conducting all kinds of assessment: quizzes, exams, standardized tests, and projects, to name just a few. Assessment is a routine task for all teachers; you may think that conducting science assessment requires no more than a commonsense approach, as Elisia and Eric, two preservice teachers think:

ERIC, ELEMENTARY PRESERVICE TEACHER

Eric enters the elementary teacher certification program with a psychology major. Among others, he is taking an elementary science methods course to learn how to teach science. The textbook for the course is a popular one commonly used by many universities and contains a chapter on science assessment. According to the course syllabus, he will have 2 weeks discussing various assessment methods. He knows assessment is an essential part of elementary science teaching because he remembers his experiences as an elementary school student in taking end-of-unit tests and state science exams and, of course, receiving a grade for science on the report card. He basically thinks assessment is a way to assess students and to give students a grade. Beyond grading students, he does not know other roles assessment may play in science teaching and leaning. As far as assessment techniques go, he feels that he has no difficulty in developing a test using multiple-choice questions, short-answer questions, and even a project. Overall, he does not think assessment should play a big role in his teaching of elementary grades, given his belief that students of such young ages need well-rounded education and development, instead of just being good at test taking. At least, he anticipates students will have diverse abilities and backgrounds, and thus no single assessment will fit them all.

(Continued)

(Continued)

Elisia, Secondary Chemistry Preservice Teacher

Elisia enters the secondary science—more specifically, chemistry—certification program with a major in chemistry. She is taking a secondary science methods course in addition to a number of other courses. Her science methods course also adopts a popular secondary science methods text in which assessment is one of its many chapters. The course syllabus shows that there will be 2 weeks of class time devoted to science assessment. Elisia thinks assessment is an essential component of her responsibility as a secondary science teacher because grades have to be given to students, and their transcripts will be used for many purposes, including applying for university admissions and scholarships. Elisia does not think conducting science assessment is a challenging task because she knows what multiple-choice, essay, project, and even standardized tests are. She is particularly familiar with multiple-choice questions because they are the dominant question type used in her high school and university science classes. Unlike Eric, Elisia believes that assessment should play a very important part of her teaching because high schools should prepare students for life, in which passing tests for various opportunities is a necessity. Thus, it is her responsibility to prepare students for various external tests.

Do Eric and Elisia's cases sound familiar to you? For them, assessment is what a teacher does *to* students. This seems to be the common conception among most preservice teachers and what they have seen routinely in science classrooms from elementary school to university. Can and should science assessment be more than grading students? Should and can science assessment be different in elementary and secondary science classrooms?

RELATIONSHIP BETWEEN ASSESSMENT AND INSTRUCTION

In thinking about how assessment should be conducted in elementary and secondary science classrooms, you first need to consider the relationship between assessment and instruction. The National Science Education Standards (NSES) Teaching Standard C states the following:

> Teachers of science engage in ongoing assessment of their teaching and student learning. In doing this, teachers
>
> a. Use multiple methods and systematically gather data about student understanding and ability.
> b. Analyze assessment data to guide teaching.
> c. Guide students in self-assessment.
> d. Use student data, observations of teaching, and interactions with colleagues to reflect on and improve teaching practice. (National Research Council [NRC], 1996, pp. 37–38)

How does the Teaching Standard C sound to you? Besides assessment of students, Teaching Standard C also suggests assessment to be used to guide and improve teaching. For many beginning science teachers, the role of science assessment in guiding and improving teaching may be unfamiliar. One key characteristic of science assessment implied in the above teaching standard is that science assessment is an integral component of science teaching and learning. Science assessment is not just one activity, such as an end-of-unit test or state exam, but also an ongoing process happening simultaneously with science teaching and learning activities. Science assessment is not only what you do to students but also what you do to inform teaching. A National Research Council committee has called for the design of a science learning environment to be assessment centered (Bransford, Brown, & Cocking, 2000), which best demonstrates the prominent role of assessment in the teaching and learning processes. Science teaching and learning in an assessment-centered learning environment are guided by assessment and, in turn, inform assessment.

Assessment is a systematic, multistep, and multifaceted process involving the collection and interpretation of data (NRC, 1996). There are four components in assessment. The *data use* component refers to the intended use of assessment results such as grading students, planning instruction, improving curricula, and comparing students. The *data collection* component refers to the target on which assessment data will be collected, such as student achievement, science inquiry ability, and attitude toward science. The *methods to collect data* refers to the specific ways to collect data, such as paper-and-pencil tests, interviews, and performance tasks. And the last component, *users of data,* refers to people or organizations that will have an interest in or make use of the data, such as students, teachers, and universities. The combination of the above four components forms assessment. Thus, assessment is a complex process; it entails systematic planning, implementation, analysis, and interpretation. The core of assessment is data, which make assessment empirical or, in other words, a scientific inquiry enterprise.

Assessment includes two processes: measurement and evaluation. **Measurement** is a process of quantifying the degree to which a student possesses a given characteristic, quality, or feature, while **evaluation** is the process of interpreting measurement data based on a set of criteria in order to make certain judgments. A key tool of measurement is a test. A **test** is a set of questions or tasks that elicit student responses plus a set of scoring keys or schemes to score them. A test can include a variety of question formats, such as multiple-choice questions, a concept mapping task, and a performance task.

Although science assessment is integral to science teaching and learning, it is also distinct from science teaching and learning. Guided by the overall purpose for supporting and improving science teaching and learning, science assessment purposefully collects relevant data, as well as analyzes and interprets them to answer specific questions about science teaching and learning. The centrality of data in science assessment requires that data to be collected are of high technical quality (i.e., meeting technical standards). Only when we are sure that assessment data are of high technical quality can we use assessment data to answer important questions about science teaching and learning. Sample questions that may be answered by science assessment are the following: What preconceptions do students bring to the science classroom? How do students' conceptions change during science teaching and learning? And have students mastered the expected learning standards?

Because science assessment and science teaching and learning are both distinct and closely related, we may consider science assessment and science teaching and learning as two sides of the same coin. Without teaching and learning, science assessment is meaningless; without assessment, science teaching and learning is mindless. Therefore, you must consider science assessment and teaching and learning at the same time when planning for effective science instruction.

APPLICATION AND SELF-REFLECTION 1.1

Give two assessment examples you experienced when you were a student, one in elementary school and another in high school, to answer the following questions:

What was the intended use of the assessment?

What was the target of the assessment and how was it measured?

How was the assessment result used?

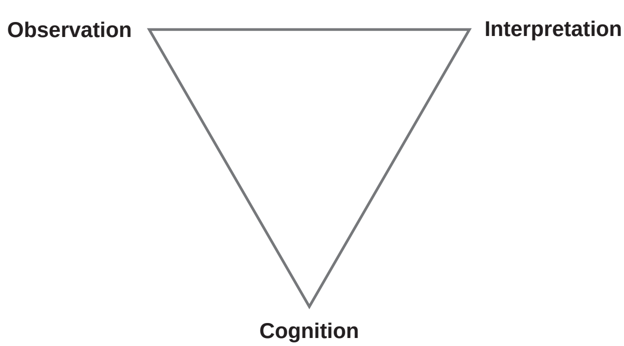
Are there matters of concern to you from your current perspective as a preservice teacher?

How should the assessment be enhanced?

FOUNDATIONS OF SCIENCE ASSESSMENT

What are fundamental considerations of science assessment? A National Research Council committee on assessment conceptualizes assessment to include three foundations:

FIGURE 1.1 The Assessment Triangle



Source: Adapted from the NRC (2001).

In Figure 1.1, **observation** refers to assessment tasks through which students' attainment of learning outcomes is elicited, **interpretation** refers to measurement models through which the assessment data are interpreted, and **cognition** refers to theories on how students learn. Observation and interpretation are related to data collection, analysis, and validation, and cognition is related to science teaching and learning. Observation and interpretation and science teaching and learning must agree with each other.

Assessment Foundation 1: Cognition

Research has suggested that the social-cultural constructivist approach to science teaching is most promising (Tobin, Tippins, & Gallard, 1994). Effective teaching follows the following principles:

- Teachers must draw out and work with the preexisting understandings that their students bring with them.
- Teachers must teach some subject matter in depth, providing many examples in which the same concept is at work and providing a firm foundation of factual knowledge.
- The teaching of metacognitive skills should be intergraded into the curriculum in a variety of subject areas. (Bransford et al., 2000)

In addition, effective teaching takes place in the following learning environments:

- Student centered: Schools and classrooms are organized around students.
- Knowledge centered: Attention is given to what is taught, why it is taught, and what competence or mastery looks like.
- Assessment centered: this consists of formative assessment—ongoing assessments designed to make students' thinking visible to both teachers and students—and summative assessment—assessments at the end of a learning unit to find out how well students have achieved the standards.
- Community centered: Develop norms for the classroom and school as well as connections to the outside world that support core learning values. (Bransford et al., 2000)

Implications of the above principles for science assessment are presented in Table 1.1.

Assessment Foundation 2: Observation

How can we collect data to fulfill the above assessment demands? Different assessment demands require different types of data to be collected. For example, techniques for identifying student preconceptions can be diagnostic inventories, interviews, and free-response writing. Similarly, many techniques are also available for assessing in-depth conceptual understanding such as two-tiered multiple-choice and constructed-response questions. Techniques for assessing application can be performance assessment and Vee diagrams.

TABLE 1.1 Cognitive Principles and Their Implications for Science Assessment

Cognitive Principles		*Assessment Demands*
1. Teachers must draw out and work with the preexisting understandings that their students bring with them.	→	1.1 Identify student preconceptions 1.2 Monitor student conceptual change
2. Teachers must teach some subject matter in depth, providing many examples in which the same concept is at work and providing a firm foundation of factual knowledge.	→	2.1 Assess in-depth conceptual understanding 2.2 Assess application 2.3 Assess factual knowledge
3. The teaching of metacognitive skills should be integrated into the curriculum in a variety of subject areas.	→	3.1 Assess metacognitive skills as part of the curriculum

Multiple-choice questions are commonly used for assessing factual knowledge. Finally, many techniques such as concept mapping and portfolio assessment are appropriate for assessing and promoting metacognition in students.

Observation may be considered as consisting of multiple dimensions. At least the following dimensions can be conceptualized:

- The medium: Observation can be based on text, audio-video, graphic, and physical action.
- The time: Observation can take place from instant responses to long-term responses ranging from days to months.
- The agent: Observation can take place individually, in pairs, and in groups.
- The construct: Observation may involve the cognitive domain (i.e., knowledge, comprehension, application, analysis, evaluation, and creation), affective domain, and psychomotor domain.
- The content: Observation may involve a single topic or multiple topics.

Assessment Foundation 3: The Measurement Model

The most commonly applied measurement model is the classical test theory (CTT). The CTT has three basic aspects: validity, reliability, and absence of bias. Validity refers to the degree to which inferences made from assessment data are accurate and sound. Assessment validation may be based on content (the alignment between assessment coverage and the intended curriculum), relevant criteria (the correlation between two sets of assessment data), and theoretical construct (the agreement with the hypothesized mental processes or products). Reliability refers to the degree to which assessment data are replicable and consistent across time, contexts, and tasks. Reliability may be established by the internal consistency of assessment items or tasks and the stability of student assessment scores over time, across contexts, or among raters. Absence of bias refers to the positive consequences of inferences and uses of assessment data. Absence of bias may be established by documenting the immediate and long-term effects of assessment result uses.

APPLICATION AND SELF-REFLECTION 1.2

Are you surprised that science assessment can be so complex and there are so many aspects to consider? Read the following two scenarios and examine how each relates to the three assessment foundations described above.

Scenario 1: Ms. A has just completed a unit on living things with her fourth-grade class. Although the textbook unit contains many colorful pictures of various living organisms and descriptions of their characteristics, she decided that a more effective way to teach this unit was to allow students to observe and explore various living organisms they encounter in their daily lives. During the unit, students were asked to select an area (i.e., a mini habitat, in their backyards) and observe and record the behaviors and characteristics of the living things in the "mini" habitats. Class time was given for students to share observations and research questions, as well as to conduct further research in the library and on the Internet. The culminating event of the unit was a poster presentation session open to the public (parents were particularly invited). Assessment was conducted during the presentation. Students were given a grade based on how well their posters were designed and how their presentations were made.

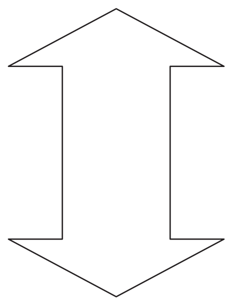
Scenario 2: Ms. P has just completed a unit on Newton's second law to her Grade 11 physics class. Ms. P's approach to teaching physics is based on a strong belief that students learn best by hands-on learning. During the unit, students spent about half the time of the unit doing labs and the other half working on textbook problems involving primarily calculations using the formula. The end-of-unit test was modeled after the state exam consisting primarily of multiple-choice questions. Student unit grades were based on the end-of-unit test scores.

SCIENCE ASSESSMENT AS A SYSTEM

What are the types of assessments you need to consider when you plan for teaching a science unit? The best way to approach planning science assessment is to consider an approach to planning for science instruction. One popular approach for instructional planning is the backward design approach elaborated by Wiggins and McTighe (2005). Briefly, the backward design approach ensures that the planned curriculum will result in real understanding instead of just factual knowledge in students. The backward design approach considers the key to developing real understanding in science as the coherence among curriculum goals, assessment evidence, and learning activities. Although the approach is described as three stages, curriculum planning does not have to follow the stages in sequence—multiple entry points are possible. The three stages are (a) Stage 1—define desired results, (b) Stage 2—identify assessment evidence, and (c) Stage 3—plan learning activities. Based on the backward design approach, a complete assessment plan supporting science teaching and learning should parallel the above three stages. Figure 1.2 shows the assessment types and their correspondence to the backward design approach stages.

FIGURE 1.2 Backward Design Approach and Its Implication for Assessment

Planning for Science Teaching and Learning		
Stage 1: Desired learning outcomes: What will students know and be able to do?	**Stage 2:** How do you know if students have obtained the desired learning outcomes?	**Stage 3:** What learning experiences and instruction will enable students to achieve the desired learning outcomes?



Diagnostic Assessment: What do students bring to the classroom and how are students' understandings different from the desired understandings?	**Summative Assessment:** What are the best ways to obtain evidence for students' mastery of the desired learning outcomes?	**Formative Assessment:** What feedback should you obtain in order to monitor students' learning trajectory toward mastering the desired learning outcomes?
Assessment to Support Science Teaching and Learning		

Based on Figure1.2, assessment may take place before, after, and during instruction, which gives rise to diagnostic assessment, summative assessment, and formative assessment. **Diagnostic assessment** is conducted specifically to identify students' strengths and weaknesses on the intended learning objectives so that effective planning for instruction may take place. **Summative assessment** takes place at the conclusion of instruction in order to grade students for mastering instructional objectives and make decisions for student future learning. **Formative assessment** takes place as part of ongoing instruction in order to monitor and make adjustment to the ongoing instruction. Obviously, the above three types of assessments are relative to the instructional context. One unit of instruction can be a part of a larger program of study, and thus the summative assessment of one unit may become a diagnostic assessment for the next one or a formative assessment for the entire program of study. Please note, in the backward design approach, evidence for student mastery of unit learning outcomes must be decided before you plan for learning activities, which is why summative assessment is planned before formative assessment.

Diagnostic Assessment

Diagnostic assessment takes place when you plan for a new sequence of science instruction, such as a new unit. The purpose is to find out what preconceptions students have related to the learning standards and how they are different from the scientifically accepted conceptions. Although research in the past three decades has documented various students' alternative conceptions related to a large number of key science concepts (Baker, 2004; Wandersee, Mintzes, & Novak, 1994), and it is likely that many of those alternative conceptions documented in the research literature apply to your students, there are a number of benefits for science teachers to conduct their own diagnostic assessment. The important benefits are to

- identify different ideas that your students specifically have,
- convey to your students that you value their ideas and take their ideas into consideration in your planning,
- motivate students to learn as the learning activities are directly relevant to their ideas,
- develop in students a sense of self-control in their learning, and
- provide a baseline for both students and you to check for learning progression (Ebenezer & Connor, 1996).

There are many ways to conduct diagnostic assessment. For example, an elementary teacher may use drawing (e.g., drawing a scientist, drawing a diagram on the relationship between a producer and consumer), brainstorming, or journaling to find students' preconceptions when planning a science unit. A secondary science teacher may use a multiple-choice or an open-ended question survey to find out what students have already known and are still confused about the concepts to be taught in a new unit. Chapter 2 will introduce a variety of diagnostic assessment methods.

Summative Assessment

Summative assessment takes place when a sequence of science instruction, such as a unit, has been completed. The purpose is to find out how well students have mastered the intended learning outcomes. Summative assessment is usually the basis for grading but should not be the only basis. Summative assessment does not have to be paper and pencil based. Depending on the desired learning outcomes, summative assessment tasks can be in various forms. The multidimensional observations discussed earlier apply here. Given the current standards-based approach to science education and accountability, external tests can also become part of the summative assessment.

Summative assessment is the most commonly used assessment activity in science teaching. In elementary grades, teachers may use portfolios, homework, in-class participation, and tests to grade students; in secondary grades, teachers may use tests, projects, and even external tests (e.g., state tests) to grade students. Chapters 3, 4, and 5 will introduce various summative assessment methods and develop specific skills related to developing summative assessments, including using standardized tests.

Formative Assessment

Formative assessment takes place as components of ongoing science instructional activities. Its purpose is to obtain feedback on students' learning in terms of the desired learning outcomes so that ongoing adjustment to the learning activities may be made. Because of the seamless integration of formative assessment into ongoing instruction, there is no need for a clear distinction between formative assessment and learning activities.

Formative assessment can be done in many different ways, both formal and informal. For example, in elementary grades, a teacher may simply use homework as a way to obtain information about the ongoing instruction and make adjustments accordingly. A secondary science teacher may periodically give students a nongrading pop quiz to find out how students are making progress during the unit. Chapter 6 will develop some important formative assessment methods and specific skills for formative assessment. Methods of grading based on both summative and formative assessment information will be introduced in Chapter 7.

It needs to be pointed out that although three types of assessment are differentiated, in reality, the distinction among them is quite fuzzy. A diagnostic assessment may also be used as formative or summative assessment, and a formative assessment may also be used as a summative assessment. Therefore, the difference among the three types of assessment is mainly conceptual; in practice, one assessment activity may perform more than one function of the assessment. No matter what assessment is used, the process must be valid, reliable, and absent of bias. Chapter 8 will introduce techniques to examine science assessment as an inquiry process by establishing its validity, reliability, and absence of bias.

Application and Self-Reflection 1.3

It should be clear to you now that planning for science assessment should be concurrent with planning for science instruction. We have differentiated three types of science assessment: diagnostic assessment, summative assessment, and formative assessment. Can you think of two assessment methods, one for an elementary science grade and another for a secondary science grade, that can be used to serve all three forms of assessment? What are the advantages and disadvantages of using one method for multiple forms of assessments?

LEGAL CONSIDERATIONS PERTAINING TO SCIENCE ASSESSMENT

When planning and conducting science assessment, teachers need to be informed of pertinent federal laws. Here we review two public laws that are relevant to all assessment situations; specific laws pertaining to specific assessments (e.g., standardized tests) will be discussed later. One is related to privacy and the other to students with disabilities. Failing to follow the laws will result in court challenges with serious consequences, including academic discipline, teaching probation, and termination of teaching certification and contract. Becoming knowledgeable about pertinent laws related to the assessment of students is a part of a science teacher's professional competence.

Assessment Information and Student Records

The Family Education Rights to Privacy Act (FERPA), enacted in 1974, guarantees that students and their parents have the right to control the use of their personal information such as student academic records. Science teachers deal with all kinds of student personal information, such as test scores, course grades, report cards, and individualized education programs for students with disabilities. Student records, whether or not on a teacher's desk or in an official student file (including computer files), are confidential and may not be released without permission. Accordingly, all assessment data must be regarded as confidential by the teacher and the school. Before a student reaches age 18, the child's parents or the guardian must give consent before any of the student's personal data are released to a third party. After a child has reached 18, the child can grant permission regarding the release of his or her personal data. In addition, the child and the parents/guardians have the right to review and challenge the accuracy of the information kept in the child's file or record. This act also implies that assessment data can be collected only by the teacher. If a third party, such as an external agency or university researcher, would like to conduct an assessment, written consent must be obtained from parents before assessment is conducted.

Specific to assessment, peer grading is a common practice in science classrooms. The Supreme Court ruling suggests that peer grading is not part of student record and thus does not violate the privacy act despite the objection to the practice by some parents (Aquila, 2008). However, if students' grades are posted in the hallway, this would clearly violate the privacy act. There are exceptions to FERPA. Information about a student that can be released includes directory information, such as the student's name, address, telephone number, and date and place of birth; major field of study; dates of attendance; participation in officially recognized activities and sports; weight and height of members of athletic teams; date of graduation; and awards received. A school may also show or turn over student records without permission to (a) other officials of the same school system; (b) certain federal, state, and local authorities performing functions authorized by law; (c) the juvenile justice system under subpoena or a court order; (d) accrediting agencies; and (e) in emergencies to protect a student's health or safety (Carin & Bass, 1997).

Test Accommodation and Alternative Assessment

Another law pertaining to assessment is the Individuals with Disabilities Education Act (IDEA). The most recent version of IDEA was passed by Congress in December 2004, replacing the 1997 version of the act. According to IDEA 2004, disability refers to mental retardation, hearing impairments (including deafness), speech or language impairments, visual impairments (including blindness), serious emotional disturbance (referred to in this title as *emotional disturbance*), orthopedic impairments, autism, traumatic brain injury, other health impairments, or specific learning disabilities. It further defines that a child aged 3 through 9 may, at the discretion of the state and the local educational agency, be considered as having disability if a child is experiencing developmental delays, as defined by the state and as measured by appropriate diagnostic instruments and procedures, in one or more of the following areas: physical development, cognitive development, communication development, social or emotional development, or adaptive development. The law regards disability a natural part of the human experience and should in no way diminish the right of students to participate in learning.

Specific to science instruction, science teachers must follow the following six principles (Carin & Bass, 1997):

- Zero reject: No student with disabilities can be excluded from a free, appropriate education.
- Nondiscriminatory evaluation: Schools must evaluate students fairly (without bias) to determine if they have a disability, identify what kind of disability that student has, and identify how extensive it is.
- Appropriate education: Schools must tailor education for individual students (individualized education program) based on nondiscriminatory education. Schools are also required to augment that education with related support services and supplementary aids.
- Least restrictive environment: Schools must educate students with disabilities alongside students without disabilities to the maximum extent appropriate for the students with disabilities. The school may not remove a student from general education unless the student cannot be educated there successfully.
- Procedural due process: IDEA provides safeguards for students against schools' actions, including a right to sue in court.
- Parental and student participation: Schools must collaborate with parents and adolescent students in designing and carrying out special education programs.

Because science assessment is closely related to science instruction, the above six principles have important implications to science assessments. There are two types of assessment activities in which science teachers may be involved: (a) initial assessment to identify a student suspected of having disability and (b) assessment accommodations and alternative assessment. Before any assessment accommodation or alternative assessment is provided, a formal evaluation of a child is needed to determine if the child has a disability. A parent of a child, a state educational agency, another state agency, or a local educational agency may initiate a request for an initial evaluation to determine if the child has a disability. Evaluation should use a variety of assessment tools and strategies to gather relevant functional, developmental, and academic information, including information provided by the parent. At least one regular classroom teacher must be a member of the evaluation team. After a child has been determined to have a disability, an individualized education program or IEP must be developed. An IEP is a written statement for each child with a disability that includes, among other things, (a) a statement of the child's present levels of academic achievement and functional performance; (b) a statement of measurable annual goals, including academic and functional goals; and (c) a description of how the child's progress toward meeting the annual goals will be measured, including any individually appropriate assessment accommodations or alternative assessments on a particular state or district-wide assessment of student achievement. It can be seen that test accommodation and alternative assessment provision are not a spontaneous decision; they are planned as part of the IEP for the child with a disability.

Various accommodations are possible depending on the nature and degree of the disability. Sample accommodations are Braille, large prints, recorded tests, individually administrated tests, oral dictation, extended time, separate rooms, and so on. Similarly, various alternative assessment forms may be provided depending on the nature and degree of disability. Sample alternative tests to paper-and-pencil tests are translation into another language, shorter version of the test, portfolio assessment, observations, performance test, and so forth.

APPLICATION AND SELF-REFLECTION 1.4

Imagine the following two cases. Decide if test accommodation, alternative assessment, or differentiated assessment is necessary for an end-of-unit test and, if yes, how and why.

Andy: a fifth-grade student identified as having a disability in reading. Andy receives science instruction together with the rest of the class. He participates in all activities, although he receives help from designated classmates or the teacher when an activity involves extensive reading. Andy does not seem to have any difficulty in understanding and mastering the science content.

Ruth: a 10th-grade student taking a biology course. She is the most advanced student in the class. It seems that she has already known and understood most of the curriculum materials; she is taking the course because this course is required by the state to graduate. Given her situation, the teacher often assigns different and more advanced homework for her to do. She is also regularly asked to help other students who are struggling in the course.

SCIENCE ASSESSMENT STANDARDS

Each of you may have different concerns related to various science assessment aspects described above. What are expectations of science teachers in terms of science assessment? The NRC (1996) developed a set of assessment standards as part of the National Science Education Standards. The NRC assessment standards contain five standards as follows:

1. Assessment Standard A: Assessment must be consistent with the decisions they are designed to inform. This standard is further elaborated into the following substandards:
   a. Assessments are designed deliberately.
   b. Assessments have explicitly stated purposes.
   c. The relationship between the decisions and the data is clear.
   d. Assessment procedures are internally consistent.

2. Assessment Standard B: Achievement and opportunity to learn science must be assessed. This standard is further elaborated into the following substandards:

   a. Achievement data collected focus on the science content that is most important for students to learn.
   b. Opportunity-to-learn data collected focus on the most powerful indicators of learning.
   c. Equal attention must be given to the assessment of opportunity to learn and to the assessment of student achievement.

3. Assessment Standard C: The technical quality of the data collected is well matched to the decisions and actions taken on the basis of their interpretation. This standard is further elaborated into the following substandards:

   a. The feature that is claimed to be measured is actually measured.
   b. An individual student's performance is similar on two or more tasks that claim to measure the same aspect of student achievement.
   c. Students have an adequate opportunity to demonstrate their achievements.
   d. Assessment tasks and methods for presenting them provide data that are sufficiently stable to lead to the same decisions if used at different times.

4. Assessment Standard D: Assessment practices must be fair. This standard is further elaborated into the following substandards:

   a. Assessment tasks must be reviewed for the use of stereotypes, for assumptions that reflect the perspectives or experiences of a particular group, for language that might be offensive to a particular group, and for other features that might distract students from the intended tasks.
   b. Large-scale assessments must use statistical techniques to identify potential bias among subgroups.
   c. Assessment tasks must be modified appropriately to accommodate the needs of students with physical disabilities, learning disabilities, or limited English proficiency.
   d. Assessment tasks must be set in a variety of contexts, be engaging to students with different interests and experiences, and must not assume the perspective or experience of a particular gender, racial, or ethnic group.

5. Assessment Standard E: The inferences made from assessments about student achievement and opportunity to learn must be sound. This standard is further elaborated into the following substandards:

   a. When making inferences from assessment data about student achievement and opportunity to learn science, explicit reference needs to be made to the assumptions on which the inferences are based.

The above science assessment standards are comprehensive and demanding; adequate and systematic preparation in knowledge and skills of science assessment is necessary. This book intends to just do that. This book will walk you through the needed assessment knowledge and skills to help you plan an effective learning unit. Those of you who are motivated to know more about current research on the assessment methods introduced in each chapter may pursue further readings suggested at the end of each chapter. Helping every teacher to meet the above assessment standards is the ultimate goal of this book.

Application and Self-Reflection 1.5

Let's take a closer look at the above assessment standards. For which standards do you think you understand the specific expectations? For which standards do you think you are not sure about the specific expectations? Place a √ to indicate "you understand" and a ? for "not sure."

Standard	*Your Feeling*
Aa. Deliberately design assessment.	
Ab. Use assessments for explicit purposes.	
Ac. Relate clearly decisions to data.	
Ad. Follow internally consistent assessment procedures.	
Ba. Collect achievement data on the science content that is most important for students to learn.	
Bb. Collect opportunity-to-learn data as the most powerful indicators of learning.	
Bc. Assess both opportunity to learn and student achievement.	
Ca. Ensure that the feature claimed to be measured is actually measured.	
Cb. Ensure that an individual student's performance is similar on two or more tasks that claim to measure the same aspect of student achievement.	
Cc. Ensure that students have adequate opportunity to demonstrate their achievements.	
Cd. Ensure that assessment tasks and methods for presenting them provide data that are sufficiently stable to lead to the same decisions if used at different times.	
Da. Review assessment tasks for the use of stereotypes, for assumptions that reflect the perspectives or experiences of a particular group, for language that might be offensive to a particular group, and for other features that might distract students from the intended tasks.	
Db. Use statistical techniques in large-scale assessments to identify potential bias among various subgroups.	

(Continued)

(Continued)

Standard	*Your Feeling*
Dc. Appropriately modify assessment tasks to accommodate needs of students with physical disabilities, learning disabilities, or limited English proficiency.	
Dd. Set assessment tasks in a variety of contexts, to engage students with different interests and experiences, and not to assume the perspective or experience of a particular gender or racial or ethnic group.	
Ea. Make sound inferences based on assessment data about student achievement and the opportunity to learn science.	

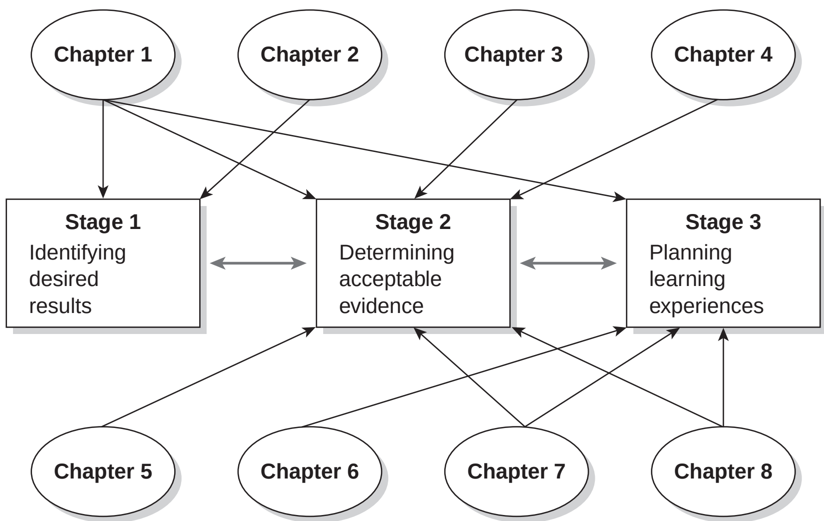
You may find yourself unsure about many of the above standards. For example, what does it mean by "deliberately design assessment"? What does it mean by "make sound inferences based on assessment data about student achievement and opportunity to learn science"? One important reason that you are not sure about these standards is because you do not yet possess enough assessment knowledge and skills. Also, you may not know specific science teaching and learning contexts in which assessment takes place. There are no absolute assessment competences without referring to specific science teaching and learning contexts. Effective science assessment must be planned and conducted within a science teaching and learning framework. This book will help you in both of the above two regards; it will introduce not only a wide variety of assessment methods but also the science teaching and learning contexts to which various assessment methods apply. Each chapter will begin with a list of pertinent assessment standards the chapter will address, followed by a list of assessment skills the chapter will help develop in order for you to meet the assessment standards. At the end of each chapter, there will also be a mastery checklist of the essential skills.

HOW THIS BOOK IS ORGANIZED

There are seven chapters following this first chapter. The seven chapters are presented in the order of diagnostic assessment (Chapter 2), summative assessment (Chapters 3, 4, and 5), and formative assessment (Chapter 6). Chapter 7 deals with grading. The final chapter, Chapter 8, deals with a common element of all assessment—data—and discusses ways of using data to improve science teaching and learning. Chapter 8 also presents a rationale of science assessment as an inquiry to integrate all chapters in this book into a coherent conceptual framework. Please note that the above order is only one way to sequence the chapters. With the exceptions of Chapters 1 and 8, which should be used as the first and last chapters, the other six chapters are relatively independent from each other and thus can be used in any order.

Guiding all chapters in this book are the assessment standards within the National Science Education Standards (NRC, 1996) and the Understanding by Design (UbD) approach (Wiggins & McTighe, 2005). At the beginning of each chapter, a box will identify the NSES assessment standards the chapter addresses, and another box will list the essential assessment skills the chapter intends to develop. Because UbD is presently a common way to structure a science methods course, relations between individual chapters and the UbD scheme are shown in Figure 1.3. Similar relations can be identified between the book chapters and other conceptual frameworks of methods courses.

FIGURE 1.3 Organization of the Book



Within each chapter, an Application and Self-Reflection will be placed at the end of each major section. At the end of each chapter, a Chapter Summary will describe major ideas and skills presented in the chapter. A Mastery Checklist is then presented for you to check if you have mastered all the essential skills. Finally, you will be directed to a Web site where you will find useful resources such as research findings, sample assessment work, Web links, assessment instruments, worksheets, and statistical analysis templates. For those who are interested in knowing more about research pertaining to the topics presented in the chapter, the Further Readings section suggests some key references to study.

The Cases of Eric and Elisia: What Is Science Assessment All About?

Eric and Elisia began this chapter without much expectation because they thought science assessment was more or less a commonsense practice. After finishing this first chapter, they now seem to realize that assessment may not be as simple as they initially thought at all. They can see how assessment is integral to instruction; thus, planning for science instruction and assessment must take place simultaneously. Furthermore, they understand that, besides grading students, assessment can play an important role in improving their teaching. They are somewhat familiar with different types of assessments (e.g., summative, formative, and diagnostic), but they did not realize how different assessment methods can form a coherent assessment system. The federal laws pertaining to assessment are new to them; they can now understand how handling student assessment information, assessment accommodation and alternative assessment provisions, and differentiated assessments can be serious and thus must be conducted within legal boundaries and with due care. The assessment standards specified in the National Science Education Standards are very comprehensive; most of them are quite vague to them. Given those assessment standards, they have now become anxious to know exactly what those standards require and how to become competent in them. Overall, this chapter seems to have opened their eyes, but they remain unclear how the following chapters will help them to become a science teacher in terms of assessment. Eric and Elisia's learning journey on science assessment continues. . . .

Do the experiences of Eric and Elisia sound familiar to your? What was your initial expectation of science assessment, and how it has changed as the result of this chapter?

Chapter Summary

- Science assessment is an integral component of science teaching and learning; it takes place within science teaching and learning contexts to support student learning. Science assessment is also distinct from science teaching and learning because one central aspect of science assessment involves data collection, analysis, and interpretation.

- Science assessment consists of three foundations: cognition, observation, and interpretation. Observation refers to assessment tasks through which students' attainment of learning outcomes is elicited, interpretation refers to measurement models through which the assessment data are interpreted, and cognition refers to theories on how students learn. Observation and interpretation are related to data collection, analysis, and validation, and cognition is related to science teaching and learning. Observation and interpretation and science teaching and learning must agree with each other.

- Planning for science assessment should take place concurrently with planning for science teaching and learning. Consistent with the backward design approach to planning for science teaching and learning, science assessment should include diagnostic assessment, summative assessment, and formative assessment.

- Applicable federal laws to assessment are the Family Education Rights to Privacy Act (FERPA) and Individuals with Disabilities Education Act (IDEA). Student assessment information is confidential and may not be released without permission. Students with disabilities have the right to assessment accommodation and alternative assessment. Assessment accommodation and alternative assessment must be planned as part of students' individualized education programs (IEPs). Students without disabilities may also need differentiated assessment. Differentiated assessment must match differentiated instruction.
- The 1996 National Science Education Standards include five science assessment standards. The science assessment standards cover purposes and uses of science assessment, scopes of science assessment, and technical qualities of science assessment.

√ Mastery Checklist

- ☐ Describe the three components of science assessment.
- ☐ Develop an assessment plan to support science teaching and learning.
- ☐ Know when test accommodation, alternative assessment, and differentiated assessment are needed.
- ☐ Know how to handle assessment information within the limit of the federal privacy act.
- ☐ State the assessment standards outlined in the National Science Education Standards.

Web-Based Student Study Site

The Companion Web site for *Essentials of Science Classroom Assessment* can be found at **www.sagepub.com/liustudy.**

The site includes a variety of materials to enhance your understanding of the chapter content. Visit the study site to complete an online self-assessment of essential knowledge and skills introduced in this chapter. The study materials also include flash cards, Web resources, and more.

Further Readings

National Research Council (NRC). (2005). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

This book presents a systematic review of current learning theories and the accompanying assessment approaches. It introduces the most promising measurement models, statistical methods, and technological advances supporting assessment. This book should be foundational reading for anyone who is interested in cognitive dimensions and psychometric models underlying educational assessment.

Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.

This book provides an excellent discussion on the backward design approach to planning instruction for understanding. Assessment is an integral component of the backward design approach. Practical planning templates and many practical examples, with some of them in science, are available in the book.

References

Aquila, F. D. (2008). *School law for K–12 educators.* Thousand Oaks, CA: Sage.

Baker, V. (2004). *Beyond appearances: Students' misconceptions about basic chemical ideas* (2nd ed.). London: Royal Society of Chemistry. http://www.chemsoc.org/pdf/LearnNet/rsc/miscon.pdf

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Carin, A. A., & Bass, J. E. (1997). *Teaching science as inquiry* (9th ed.). Upper Saddle River, NJ: Merrill.

Ebenezer, J. V., & Connor, S. (1996). *Learning to teach science: A model for the 21st century*. Upper Saddle River, NJ: Merrill.

National Research Council (NRC). (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council (NRC). (2001). *Knowing what students know: The science and design of educational assessment.* Committee on the Foundations of Assessment. Washington, DC: National Academy Press.

Tobin, K., Tippins, D. J., & Gallard, A. J. (1994). Research on instructional strategies for teaching science. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 45–93). New York: Macmillan.

Wandersee, J. H., Mintzes, J. J., & Novak, J. D. (1994). Research on alternative conceptions in science. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 177–210). New York: Macmillan.

Wiggins, G., & McTighe, J. (2005). *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.