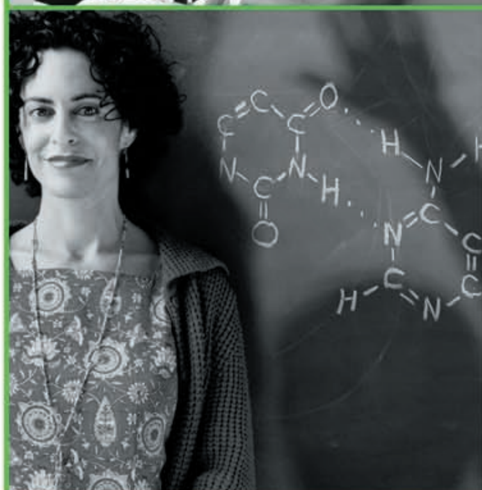
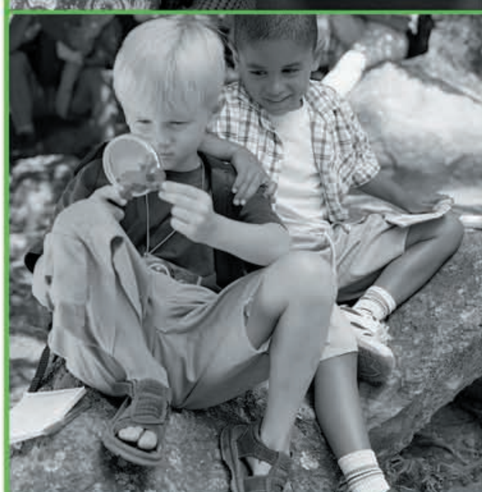
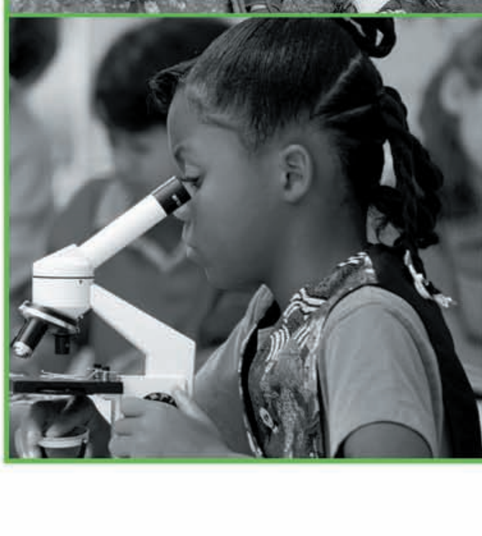


How to Assess Student Performance in Science

Using Classroom Assessments to Enhance Learning




SERVE Center
at the University of North Carolina
at Greensboro

How to Assess Student Performance in Science:

Using Classroom Assessments to Enhance Learning

First Edition, 2005

SERVE Center
at the University of North Carolina at Greensboro

Associated with the School of Education, University of North Carolina at Greensboro

PRODUCED BY

The SERVE Center at the University of North Carolina at Greensboro

WRITTEN BY

Susan M. Butler, Ph.D., Senior Program Specialist, SERVE
Nancy D. McMunn, Project Director, SERVE

EDITED BY

Christy Casbon, Editorial Consultant
Donna Nalley, Ph.D., Director of Publications, SERVE

DESIGNED BY

Tracy Hamilton, Assistant Program Specialist, SERVE

This document was originally developed with funding from the Institute of Education Sciences, U.S. Department of Education, under contract no. ED-01-CO-0015.

Table of Contents

Acknowledgments

This manual builds upon the R&D experience acknowledged in Volume I, "How to Assess Student Performance in Science: Going Beyond Multiple-Choice Tests." We gratefully acknowledge the contributions of the teachers involved in this project.

In addition, this manual was reviewed by both assessment and science experts. These reviewers provided many valuable suggestions that enhanced the quality of this publication. These reviewers included:

- LaShanda Colbert, Science Specialist, Mississippi Department of Education, Jackson, Mississippi
- Patricia Schenck, Resource Teacher, Assessment and Accountability Division, Bay District Schools, Panama City, Florida
- Wilma Guthrie, National Board Certified Science Teacher, Elementary level, Hanceville, Alabama
- Keith Eades, Assistant Superintendent, Currituck County Schools, Currituck, North Carolina
- Susan M. Brookhart, Educational Consultant, Duquesne University and Brookhart Enterprises LLC, Helena, Montana
- Dawn Hudson, Science/Health Specialist, National Board Facilitator, Paulding County Schools, Dallas, Georgia
- Michael Vigliano, Math and Science, SERVE, Greensboro, North Carolina
- Patricia McClure, Math and Science, SERVE, Greensboro North Carolina

Current Views on Assessment

Purpose of This Manual

The purpose of this manual is to enhance the use of formative assessment practices in the classroom. These formative practices are ones that support students *while they are learning*. The practices also ensure that students receive feedback and guidance on their performances in a timely manner so they can make improvements to those performances before being evaluated. In the pursuit of enhancing formative assessment practices, this publication is designed to help teachers do the following:

1) Clarify science learning targets for their students (**CHAPTER TWO**).
2) Understand the range of assessment methods available to teachers and explore the alignment of these methods to learning targets (**CHAPTERS THREE** and **FOUR**).
3) Analyze assessment information to determine individual student learning needs (**CHAPTER FIVE**).
4) Modify instructional and/or assessment practices to enhance student learning (**CHAPTER FIVE**).
5) Reflect on current grading practices and determine if current methods accurately portray student achievement (**APPENDIX**).

This publication is not intended simply as an informative text but also as a resource that encourages self-reflection on current practices. We hope you will interact with it, respond to the questions posed, implement some of the suggested strategies, and then reflect on such implementation. Through these activities, this manual can serve as an opportunity for you to examine your science assessment practices. It is most helpful if two or more teachers work through the manual together. The synergy that results from professionals sharing ideas will greatly enhance the learning experience.

Introduction

Before we can concentrate on "how to assess," we must first define *assessment*. What does it mean to "assess"? What is the purpose of assessment? Are there different methods of assessment? With a clear definition of assessment and a definite purpose given for this activity, we can then begin to develop assessments appropriate for science classes.

Assessment simply means a process, strategy, or tool for collecting information about the learning process of individuals or groups. Such assessment has two broad goals: 1) to provide feedback to students about their own learning, and 2) to monitor the effectiveness of the instruction. A variety of assessments may be used to monitor the effectiveness of instruction, ranging from the large-scale achievement tests used to calculate Adequate Yearly Progress under the No Child Left Behind Act down to the observation of students' facial expressions by a lecturing teacher. Students can also receive feedback in diverse ways (e.g., comments on a lab report or scores on a high-stakes test).

In this publication, however, we will concentrate on assessments that occur on a day-to-day basis within the science classroom. Such *classroom assessments* constitute an ongoing process in which teachers and students interact to promote enhanced student learning. This process involves using a variety of assessment strategies and tools to collect information about student learning and then using the data collected to diagnose learning problems, monitor student progress, or provide meaningful and timely feedback to students. Therefore, the main purpose of this publication is to promote an assessment process that enhances student achievement in science.

This purpose of using classroom assessment to improve student performance is echoed throughout the current classroom assessment research literature. Some (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998) label such assessments as *formative* in nature and emphasize that assessment is formative "when the evidence is actually used to adapt the teaching to meet student needs" (Black & Wiliam, 1998, p. 140). Chappuis, Stiggins, Arter, & Chappuis (2004, p. 35), echoing a term coined by Black and Wiliam, characterize this enhancing of student performance assessment process as "assessment **for** learning" and state that when "they assess for learning, teachers use classroom assessment and the continuous flow of information about student achievement that it provides to advance, not merely check on, student learning." Such assessment for learning is assessment that helps students identify the strengths and weaknesses of their performance so that they can improve their achievement. It is differentiated from "assessment **of** learning," which simply provides a means of rating students, or comparing them to one another. Assessment of learning, unlike assessment for learning, does not focus on feedback for improvement.

What is the Classroom Assessment Cycle?

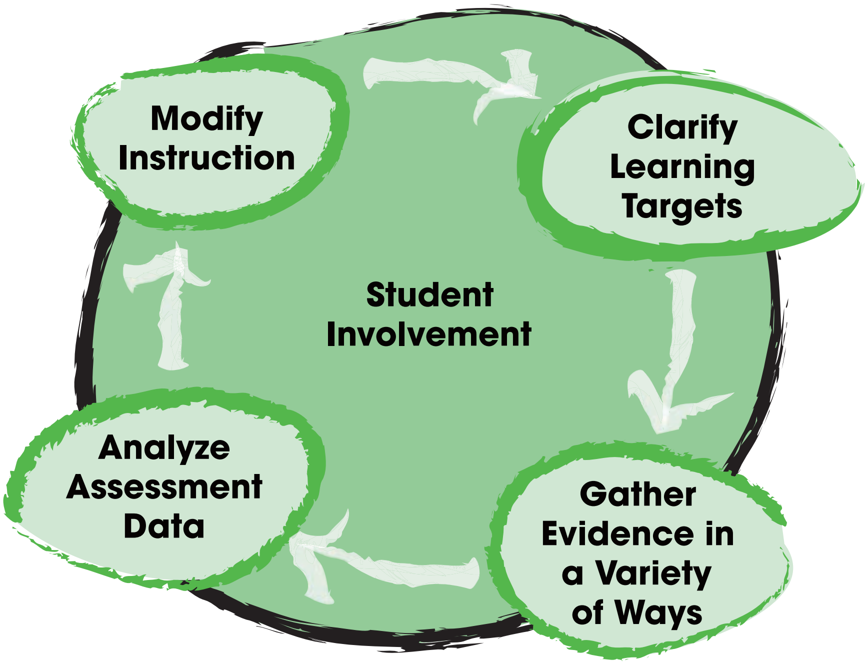
In order to fulfill the goals listed above and to promote assessment for learning, we must first explain the Classroom Assessment Cycle. This cycle outlines an assessment process that focuses on improving student performance. The four main steps of the cycle are:

1) Clarifying learning targets
2) Gathering evidence in a variety of ways
3) Analyzing assessment data
4) Modifying instruction

Let us examine a common assessment scenario and analyze it to see how it conforms to each step in the Classroom Assessment Cycle.

> **SCENARIO ONE:** Mr. Wei, a sixth-grade teacher, teaches a unit on soil formation and then gives a unit test with multiple-choice, true/false, and matching items to assess students' retention of the information. Students are told about the test one week in advance and are instructed not to bring resource materials with them to the test. Students' tests are scored and returned and form the basis of the six weeks' grade. The teacher moves on to a new unit.

FIGURE 1.1 Classroom Assessment Cycle



CLARIFYING LEARNING TARGETS: There is no mention in the above scenario of learning targets. What is important for the students to know about soil formation? Are the important concepts drawn from the sixth-grade curriculum? Does the teacher expect the students to simply learn factual information, or are other concepts important? For instance, did they learn skills in operating scientific equipment, critical thinking skills, skills in creating such products as lab reports, papers, models, etc.? For the scenario to exemplify the Classroom Assessment Cycle, the teacher would need to set particular learning targets for the instruction, plan the assessments and the instruction based upon these targets, and most importantly, *share these learning targets with students* prior to the actual instruction.

GATHERING EVIDENCE IN A VARIETY OF WAYS: The scenario mentions only one assessment—the unit test. It does not mention any formative assessments that may have prepared students to take this test. For example, did the teacher check comprehension of concepts before scheduling the test? How was this done? What day-to-day practices helped assure the teacher that students were ready for the unit test? Since this test is the only assessment mentioned and because it primarily uses only one form of assessment (that of the selected response), this scenario does not support the idea of gathering evidence in a variety of ways.

Selected responses are those in which students choose their answers from a given list, as with multiple-choice, true/false, or matching questions. Mr. Wei's test did not incorporate opportunities for students to construct their own answers. Student-created essays, short-answers, graphs, or illustrations would all be examples of *constructed responses* that could provide a window into student thinking and reveal student understanding of the concepts being tested.

ANALYZING ASSESSMENT DATA: The teacher apparently did not collect any data to determine that students were prepared for the test, and he also failed to provide opportunities for students to receive feedback on their performances before being tested. The testing data itself does not appear to be analyzed, as the teacher immediately moves on to another unit. To comply with this quadrant of the Classroom Assessment Cycle, the teacher would need to use the assessment data to make inferences about student learning, focusing not only on the whole class, but on individual student learning needs. The assessment data would need to be collected prior to instruction, during instruction, and after instruction. Mr. Wei appears to only collect data after instruction, and even this data he does not appear to analyze.

MODIFYING INSTRUCTION: Since the assessment data from the ending exam were not analyzed, no instructional modifications were made. From the scenario, it appears that the teacher will simply move on to a new unit, regardless of the assessment results.

If Scenario One does not comply with the precepts of the Classroom Assessment Cycle, what might such compliance look like? Scenario Two provides a classroom exemplar of this cycle in action.

> **SCENARIO TWO:** Mr. Rahib, a high-school chemistry teacher in Florida, is planning a unit on the Periodic Table of Elements. Mr. Rahib wants his students to understand the organization of this table and be able to use the table to predict properties of particular elements. Before planning the unit, he consults the state standards and chooses two that are appropriate for the unit, including "The students will know that elements are arranged into groups and families based on similarities in electron structure and that their physical and chemical properties can be predicted" and "The students will know that investigations are conducted to explore new phenomena, to check on previous results, to test how well a theory predicts, and to compare different theories." Mr. Rahib decides that the culminating assessment will be a performance-based one in which students determine the physical and chemical properties of particular elements and then predict the physical and chemical properties of untested elements. To prepare students for this assessment, Mr. Rahib decides to: a) assign textbook readings, b) hold class discussions of these readings, c) give a lecture on the atomic theory, d) provide students with a set of cards showing physical characteristics of persons and then help them organize the cards into "families" and predict how missing family members may look, e) ask students to provide explanations of the atomic theory and of the organization of the periodic table in their science journals, and f) assign a laboratory investigation of physical and chemical properties of particular elements, with a concomitant laboratory report. He then designs rubrics (scoring guides) to help him assess the class discussions, science journals, and lab reports, as well as one for the culminating assessment. He shares the learning targets with students, using language they understand, before beginning the unit. As students progress through the unit, they are provided with written feedback via the rubrics as

well as verbal feedback from Mr. Rahib. Mr. Rahib uses these formative assessments (a–f, above) to align his instruction with student learning needs and to determine when students are adequately prepared for the culminating assessment. He uses both the formative and the culminating (summative) assessment results to determine which students have difficulties in making predictions and selects instructional activities that will aid these students. These may include one-on-one instruction, peer tutoring, use of science manipulatives, etc.

In comparing the two scenarios, we see that the students in Mr. Rahib's class are provided with the learning targets prior to instruction, they have multiple opportunities to receive feedback on their learning, and the assessments utilized are varied and diverse. Mr. Rahib constantly monitors student learning via scoring rubrics, analyzes the assessment data to determine if further instruction is necessary, and then modifies instruction accordingly, both for the entire class and for individual students. In Mr. Rahib's class, it is also evident that students are expected to do more than simply learn facts. This did not appear to be the case in Mr. Wei's class.

Traditionally, the goal of most subject area courses (like Mr. Wei's science class) has been for students to be able to recognize or recall important facts, concepts, or relationships that have been explained to them. In the past, we have almost exclusively valued students' success at retaining and bringing forth a sample of the information they have retained. When a teacher emphasizes factual knowledge on a test, students conclude that remembering facts is the goal. When students are not given an opportunity to retest or improve their work, they may conclude that improvement is not valued. If higher-order thinking, problem-solving, and critical thinking are valued, then classroom assessments should lend value to them. It is important to remember that how and what we test sends a clear message to students about what is valued.

A key point to remember as you go through this manual is that assessing students involves gathering information about what they know or can do. If throughout 12 years of school students are assessed only on passive, non-creative work (worksheets, multiple-choice tests), how likely is it that they will become problem-solvers, creative producers, effective communicators, and self-directed learners? In Chapter Two, we will explore a variety of learning targets for students that will enable them to reach their full potentials.

APPLICATION

You may wish to reflect upon your own current assessment practices to determine if you are using all the quadrants in the Classroom Assessment Cycle with your science classes. Figure 1.2 may be of help in this process.

FIGURE 1.2
Reflecting on Your Assessment Practices

Directions: *Place a check in the column that best matches your current practices.*

STATEMENT	OFTEN	SOMETIMES	RARELY	NEVER
1. I use the state curriculum to help identify student learning targets.				
2. I share the learning targets with students before beginning instruction.				
3. I emphasize many types of learning targets in my teaching, not relying on factual learning targets alone.				
4. I plan my assessments before planning my instruction.				
5. I incorporate multiple formative assessments within each unit, so students have many opportunities to receive feedback on their learning.				
6. I create rubrics (scoring guides) to help students understand their academic weaknesses and strengths.				
7. I examine assessment data and make inferences about whole class and individual student learning.				
8. I use formative assessment data to revise/ retool my teaching within the unit.				
9. I use formative assessment data to individualize instruction for struggling students.				

To aid in identifying areas on which you may wish to concentrate, consider the following:

STATEMENTS 1–3 relate to the "Clarify Learning Targets" quadrant of the Classroom Assessment Cycle.

STATEMENTS 4–6 relate to the "Gather Evidence in a Variety of Ways" quadrant of the Classroom Assessment Cycle.

STATEMENTS 7 relates to the "Analyze Assessment Data" quadrant of the Classroom Assessment Cycle.

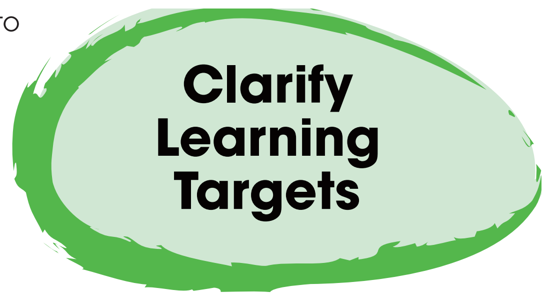
STATEMENTS 8–9 relate to the "Modify Instruction" quadrant of the Classroom Assessment Cycle.

Clarifying Learning Targets

CHAPTER TWO

What Do We Want Students to Know and Be Able to Do?

Before we can assess and teach our students, we must first decide what we want them to know and be able to do. We usually start with broad goals and then break those down into smaller components. For example, *The National Science Education Standards* publication (1996, p. ix), written by the National Research Council, begins with this goal statement: "This nation has established as a goal that all students should achieve scientific literacy." The booklet goes on to describe literacy: "Scientific literacy enables people to use scientific principles and processes in making personal decisions and to participate in discussions of scientific issues that affect society" (1996, p. ix).



With this description, the National Research Council begins to break its overall goal (scientific literacy) into smaller component parts, as depicted in **FIGURE 2.1**. One component emphasizes scientific knowledge of concepts/principles, while the other emphasizes knowledge of scientific processes.

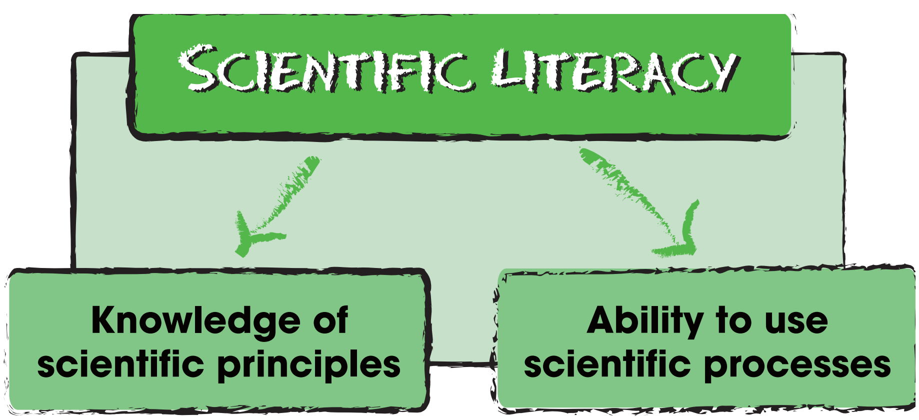


FIGURE 2.1

To aid students in acquiring knowledge of scientific principles and processes, the National Research Council has written content standards within eight different categories:

- Unifying concepts and processes in science
- Science as inquiry
- Physical science

- Life science
- Earth and space science
- Science and technology
- Science in personal and social perspectives
- History and nature of science (1996, p. 104)

Within each of these eight categories, we find the actual content standards. Each science content standard is then further explicated by a "Guide to the Content Standard" section. **FIGURE 2.2** displays this organizational scheme graphically using examples from the K–4 level.

Category: Science as Inquiry

Content Standard A:

As a result of activities in grades K–4, all students should develop:

- Abilities necessary to do scientific inquiry
- Understanding about scientific inquiry (1996, p. 121)

The Guide to the Content Standard, "abilities necessary to do scientific inquiry," includes the following tasks that students will be expected to accomplish:

- Ask a question about objects, organisms, and events in the environment
- Plan and conduct a simple investigation
- Employ simple equipment and tools to gather data and extend the senses
- Use data to construct a reasonable explanation
- Communicate investigations and explanations (1996, p. 122)

FIGURE 2.2

In this manner, the National Science Standards are broken down from complex, generalized statements to more specific ones. The tasks listed in **FIGURE 2.2** under the "Guide to the Content Standard" category show the most specific level included in the National Science Standards. The implication is that if students can complete all these tasks successfully, then they have the "abilities necessary to do scientific inquiry."

Like the National Science Education Standards, your own state curriculum may begin with broad goals that are then dissected into increasingly smaller component parts. For example, in Florida, the state science curriculum begins with the subject area (science), then breaks this into strands (The Nature of Matter, Energy, Force and Motion, Processes That Shape the Earth, Earth and Space, Processes of Life, and How Living Things Interact With Their Environment). Under each of these strands, a number of standards further explain what students should know and be able to do. The standards are then supported by benchmarks, which occur as the most specific level of the curricular hierarchy. Let's take a look at examples of these different levels from the K–2 science curriculum displayed in **FIGURE 2.3**. (Sunshine State Standards, 2005, **www.firn.edu/doe/curric/prek12/pdf/prk-2g.pdf**).

CURRICULAR LEVELS IN FLORIDA'S SUNSHINE STATE STANDARDS

FIGURE 2.3

GRADE LEVEL: K-2

SUBJECT AREA: Science

STRAND: Nature of Science

STANDARD: The student understands that all matter has observable, measurable properties.

BENCHMARKS:

1. Knows that objects can be described, classified and compared by their composition (e.g., wood or metal) and their physical properties (e.g., color, size, and shape).
2. Recognizes that the same material can exist in different states
3. Verifies that things can be done to materials to change some of their physical properties (e.g., cutting, heating, and freezing), but not all materials respond the same way (e.g., heating causes water to boil and sugar to melt)

In this figure, we note the same trend we observed in the National Science Education Standards. Each set begins with more general goals and moves to higher specificity.

APPLICATION

Examine your state science curriculum. Do you find this same trend—from generalities to specificities—within your curriculum?

Of course, greater specificity makes it easier for the teacher to determine what students are required to know and be able to do. Such specificity enables teachers to plan meaningful instruction as they know what concepts students will need to understand and what abilities students will need to develop. One of the authors of this publication once experienced some confusion with a state science standard that read, "Students will appreciate the diversity of life." Because of the general nature of this statement and the huge scope of scientific concepts this standard could encompass, it was hard to know exactly what content knowledge was being required of the students! More specificity, as is now the trend in state and national curricula, certainly makes written goals easier to interpret and implement.

Content standards and benchmarks, then, help teachers understand what students need to know and be able to do. Teachers use this knowledge to plan meaningful instruction that will prepare students for assessments based upon these standards and benchmarks. However, are standards and benchmarks truly specific enough? Will teachers completely understand how they should assess and how they should teach by simply looking at the standards and benchmarks for their science courses? We have often found that more clarification is needed for teachers to truly understand what they are required to teach. Therefore, this next section of the manual describes ways to "unpack" the standards and benchmarks so that the intent of each can be clearly understood.

Learning Targets

Some professional developers believe categorizing standards and benchmarks into *learning targets* can aid in this unpacking process (Stiggins, 2001; Regional Educational Laboratories, 1998; Butler & McMunn, 2006). Stiggins (2001, p. 66) describes five categories of such learning targets:

- Knowledge
- Reasoning
- Performance Skills
- Products
- Disposition

Let's take some state science competencies (standards) and objectives (benchmarks) from North Carolina's science curriculum and see how they can be further classified into learning target categories. (See **FIGURE 2.4**.)

Examples of Learning Targets Found in North Carolina Science Standards and Benchmarks

FIGURE 2.4

Knowledge

6TH GRADE: Explain the model for the interior of the earth.

6TH GRADE: Describe the flow of energy and matter in natural systems.

Reasoning

6TH GRADE: Analyze evidence to explain observation, make inferences and predictions, develop the relationship between evidence and explanation.

6TH GRADE: Evaluate technological designs for application of scientific principles, risks & benefits, constraints of design, consistent testing protocols.

Performance Skills

6TH GRADE: Use oral and written language to communicate findings, defend conclusions of scientific investigations.

9TH GRADE: Design and conduct scientific investigations to answer biological questions.

Products

3RD GRADE: Using shadows, follow and record the apparent movement of the sun in the sky during the day.

6TH GRADE: Prepare models and/or computer simulations to test hypotheses, evaluate how data fit.

Source: North Carolina Dept. of Public Instruction, Science Standard Course of Study and Grade Level Competencies, www.dpi.state.nc.us/curriculum/science/standard/toc.html

KNOWLEDGE TARGETS are analogous to the lower levels of Bloom's Taxonomy (Recall and Comprehension). You might think of knowledge targets as "knowing the facts." You can recognize standards and benchmarks that would be classified as knowledge targets because they often start with a few key verbs: *identify*, *list*, *explain*, *describe*, *know*. Knowledge targets are important because all of the higher-level thinking skills depend upon students possessing a knowledge base. Therefore, when we ask students to learn such laws/principles as, "For every action, there is an equal and opposite reaction" or "Matter is neither created nor destroyed," we are using knowledge targets. Knowledge targets also encompass such vocabulary definitions as, "matter," "photosynthesis," and "metamorphosis."

REASONING TARGETS abound in science classes. The key verbs found in standards and benchmarks that denote such reasoning targets are: *classify*, *categorize*, *compare*, *contrast*, *analyze*, *evaluate*. When we ask students to differentiate between eukaryotes and prokaryotes or to classify elements as solids, liquids, or gases at room temperatures, we are using reasoning targets. Writing conclusions to lab reports of scientific investigations will also call for reasoning as students must reflect on their observations, look for patterns, form inferences, and then write rationales for these inferences. This analysis process is definitely a reasoning one. Problem-solving activities also fall within the reasoning targets category, especially when students must evaluate several alternative solutions and choose the "best" one.

Students learning to operate scientific equipment, make oral presentations, use graphing calculators or use computer applications are all utilizing **PERFORMANCE SKILLS TARGETS**. Teachers can recognize standards and benchmarks that call for performance skills when they see such key verbs as: *demonstrate*, *show*, *operate*, *use*, *assemble*, *conduct*. In science class, students are called upon to learn many different types of skills. These may range from such abstract skills as comprehending subject-specific text to more tangible, concrete skills as focusing a microscope. Each skill is important and will aid students in becoming more scientifically literate and adept.

Besides learning skills in science class, students also are required to construct products. Students may keep science journals, write lab reports, construct models, create illustrations, make posters, etc. All of these would qualify as products and would fall under the **PRODUCTS LEARNING TARGET** category. Key verbs to look for in standards and benchmarks for product learning targets are: *make*, *construct*, *depict*, *write*, *create*, *design*, *record*.

DISPOSITIONAL TARGETS encompass the "unwritten" expectations we have for students. Some of these expectations may include:

- Be self-motivated.
- Exhibit positive attitudes.
- Work cooperatively.
- View self as capable.

Such dispositional learning targets do not usually appear in the written objectives or benchmarks for an academic subject. These targets are not usually graded; it is more common for them to be assessed

RESOURCE

Sample dispositional targets used in Georgia may be viewed at **www.glc.k12.ga.us**. These targets are termed "Character Education Objectives" and are required to be taught in all K–12 subject areas.

informally. They are also more susceptible than the other achievement targets to "teacher-pleasing" or socially acceptable responses. Dispositional targets are rarely entirely cognitive in nature; they are more motivational and affective. Just because these targets do not appear in curricula, however, does not mean that teachers do not have such goals in mind for students.

Why Should We Classify Standards and Benchmarks Into Learning Target Categories?

We have examined some science standards and benchmarks and then classified them into learning target categories. Why is this important or necessary? The primary reason for doing so is to clarify, as specifically as possible, what we want students to know and be able to do. We need to understand if the standard or benchmark is asking for students to learn a skill, use reasoning or critical thinking, know basic facts, or create a quality product.

We have already stated that increased specificity helps the teacher understand the learning requirements for students. The use of learning targets simply enhances the specificity of the standard or benchmark. When we examine standards and benchmarks for embedded learning targets, we are "unpacking" these statements and teasing out the "meat" or substance from them. We are using the learning targets to help us understand exactly what students should know and be able to do. In the next chapter, we will learn that knowing the learning targets can aid us in creating valid assessments—assessments that truly test whether a student has met a particular target.

Unpacking the Curriculum for a Unit of Study

Before we move to the next chapter, let's take a look at a fifth-grade unit of study to see how learning targets may overlap and interconnect to teach science concepts. This unit of study covers earth and space science concepts. The teacher, Mrs. Vega, consults the state standards for Mississippi and finds the following standard and benchmarks (Mississippi State Curriculum Frameworks, **http://marcopolo.mde.k12.ms.us/frameworks/science/sci_05.html**):

6. Investigate the structure of the Earth.
    a. Investigate the structure of the atmosphere (gas-air), hydrosphere (liquid-water), and lithosphere (solid-land).
    b. Examine how organisms affect the composition of the Earth and its atmosphere.
    c. Analyze processes that cause changes on Earth.
    d. Explore fossils as indicators of how life and environmental conditions have changed.

FIGURE 2.5 displays a unit planning sheet Mrs. Vega developed for this Earth Science unit. In this planning sheet, Mrs. Vega unpacks the curriculum into embedded learning targets and then begins to list assessments that the students will do to prove they successfully met each learning target. She then brainstorms possible instructional activities that will prepare students for these assessments. By creating such a planning sheet, Mrs. Vega ensures that she actually addresses each of the learning targets embedded in her curriculum.

FIGURE 2.5
Unit Planning Sheet

LEARNING TARGETS	SOURCE IN CURRICULUM	POSSIBLE ASSESSMENTS	POSSIBLE INSTRUCTIONAL ACTIVITIES
KNOWLEDGE	**6a–d:** All require some knowledge, as vocabulary definitions, components of the Earth and atmosphere, knowledge of processes	• KWL sheet • Vocabulary worksheet: Fill-in-the-blanks • Vocabulary quiz: Matching • Solid, Liquid, Gas worksheet: Label and Matching • Journal entries	• Textbook readings • Teacher lecture • Comprehension questions • Vocabulary wall • Demonstrations • Modeling • Class discussions • Pair shares
REASONING	**6b:** Analysis of effects **6c:** Analysis of processes **6d:** Analysis of how fossils are indicators and how life has changed	• Short-answer questions • Concept-map rubric • Higher-order questions • Cause/effect tables	• Brainstorming • Concept mapping • Demonstrations • Inquiry-based experiments • Seminars/class discussions • Before/after pictures
PERFORMANCE SKILLS	**6a:** Must learn skills needed to conduct an investigation (process skills)	• Observation of Skills checklist • Writing hypotheses worksheet	• Demonstrations • Inquiry-based experiments • Set class a problem to solve—have students note steps they took to solve it
PRODUCTS	**6b:** Composition of Earth model **6d:** Fossil experiment lab report	• Lab Report Rubric • Practice lab report • Earth Composition Model Rubric	• Work together to create rubrics • Critique products from previous years
DISPOSITIONS	None in curriculum, but want students to be able to work cooperatively in groups	• Teacher observations • Notify rule-breakers of particular rule broken	• Brainstorm rules for group work • Prioritize rules

This planning sheet provides Mrs. Vega with a good basic unit plan. She now needs to create a pacing guide (e.g., How many days will the unit take? What will be done each day? Which particular benchmarks will be covered each day?) and choose the specific activities (e.g., What experiments will the students do? How will inquiry activities be implemented? What readings do they need to complete? What problem will be used to introduce performance skills in investigations?) and design the assessments (e.g., What short-answer questions will be asked? What criteria will be used to judge the model of the Earth's composition?) Once these questions are answered, Mrs. Vega can create her daily lesson plans feeling confident that these plans include instruction and assessments aligned with the learning targets found in the curriculum.

As an example of moving from the unit plan to a daily lesson plan, let's look at the reasoning target in **FIGURE 2.5**. Mrs. Vega felt that reasoning was embedded in three of the four benchmarks. The key words she observed in these were "examine," "analyze," and "explore." In all three benchmarks, there appeared to be cause-and-effect relationships, as students were asked to examine how organisms affected the composition of the Earth, to analyze processes that caused changes on the Earth, and to explore how fossils served as indicators of change. To successfully complete these benchmarks, students will need to understand cause and effect and be able to explore relationships. Mrs. Vega is going to use brainstorming activities, graphic organizers, class discussions, and some inquiry activities (hands-on experimentation, student-designed labs) to help students make connections between cause and effect. For example, she might begin with benchmark 6c by showing pictures of the rock formations in Monument Valley, Utah, and asking students to brainstorm how the monuments were formed. Through a class sharing/discussion session, the word "erosion" may be introduced to the class. Once this word appears, Mrs. Vega may then ask how erosion occurs (e.g., What are the agents of erosion?). Once such agents as wind, water, and frost have been identified by the students, Mrs. Vega may ask students to draw a "before" picture for one of the Monument Valley photographs, identifying the causal agent for the change. Similarly, the class may examine "before" and "after" pictures of Mt. St. Helens. Here, Mrs. Vega may ask the cause of the change seen. Through such demonstrations and class discussions, Mrs. Vega is introducing cause and effect. Later, Mrs. Vega may ask students to manipulate paper or stacks of colored towels to simulate mountain building. Again, she will draw students' attention to the cause (force applied) and the effect (mountain building). She may conclude this lesson on cause and effect by asking students to create a concept map around the phrase "forces that shape the Earth."

ONE FINAL NOTE: ***Learning targets should be shared with students prior to the actual instruction or assessment of that learning target.*** Knowing the target is necessary for successful learning to occur. Think back on your own experiences as a student. Was there a time when you did not understand what you were supposed to learn? Was this a frustrating experience for you? Or, have you ever misjudged the teacher's expectations? When did you find out you had made this misjudgment? Was it after you received a grade on an assignment? Could you have done better on the assignment if you had simply understood what you were supposed to learn?

When we share the learning targets with students upfront, we are making it possible for them to hit the target. When we fail to share the learning targets, we are asking students to shoot their arrows at an invisible (or perhaps moving) target.

APPLICATION

You may practice unpacking standards and benchmarks into component learning targets by completing the matching task in Figure 2.6.

OR

You may create a unit planning sheet similar to the one in Figure 2.5 for your next unit.

FIGURE 2.6
Practice in Classifying Learning Targets

Directions: Match the standards or benchmarks on the right side to the correct learning target category.

LEARNING TARGET CATEGORY	STANDARD OR BENCHMARK
a. Knowledge b. Reasoning c. Performance Skills d. Products e. Dispositions	1. Create a model of the cell and label all parts. 2. Persevere to overcome roadblocks. 3. Identify the parts of a microscope. 4. Make a hypothesis. 5. Analyze the effects on solubility by changing the temperature of the solution. 6. Demonstrate the operation of a Bunsen burner using correct safety procedures. 7. Work cooperatively in lab, sharing equipment equitably. 8. List characteristics of living things. 9. Write a lab report. 10. Share scientific findings with the class in an oral report.

Answers: *1.d, 2.e, 3.a., 4.b, 5.b, 6.c, 7.e, 8.a, 9. d, 10.c*

CHAPTER THREE

Gathering Evidence in a Variety of Ways

METHODS OF ASSESSMENT

In Chapter Two, we explored documents that help teachers determine what students should know and be able to do. Specifically, we looked at the National Science Education Standards and at some examples of state standards. We noted a trend for standards to start as general statements and become increasingly specific as they are broken into objectives or benchmarks. Then, we experimented with unpacking such statements into component learning targets: Knowledge, Reasoning, Performance Skills, Products, and Dispositions. In this chapter, we will emphasize aligning our assessments with these learning targets, in order to create valid assessments. Assessments are *valid* when they actually provide information about student achievement of the intended learning target. An example of an invalid assessment of the ability to use a microscope correctly (a Performance Skills learning target) would be to give a pencil-and-paper test on the parts of the microscope. A more valid assessment would be to hand the student a slide and have him focus this under low and high power. Before we can begin creating these valid assessments, we need to look at the different methods of assessment available for teachers to use.

Gather Evidence in a Variety of Ways

Methods of Assessment

FIGURE 3.1 graphically displays several different methods of assessment that we will reference in this chapter. As the table shows, one way that methods of assessment can be categorized involves the use of two main headings: *Selected Response* and *Constructed Response*. Selected response assessments provide students with a list of possible answers. The students then select an appropriate answer from this list. Examples of Selected Response assessments include multiple choice questions, true/false questions, matching questions, and sometimes fill-in-the-blank questions (if students are given a word bank of answers from which to choose their response). Constructed Response assessments, unlike selected response, require students to create their own answers. **FIGURE 3.1** shows that constructed responses can be further sub-divided into *Products* and *Performances*. Short-answer questions and essay questions would fall under the Constructed Response-Product subcategory. Here, students are asked to construct their own responses to the question and to create a written record (a product). Other products listed here include science logs or journals, research papers, student-constructed graphs, portfolios, notebooks, and such graphical organizers as flow charts or concept maps.

FIGURE 3.1
Assessment Methods and Approaches

SELECTED RESPONSES	CONSTRUCTED RESPONSES	
	Products	**Performances**
Multiple Choice True/False Matching Fill-in-the-Blank	Essays Logs Journals Graphing Portfolios Paragraphs (short answers) Notebooks Flow Charts Concept Maps Research Papers	Laboratory Practical Exam Oral Presentations Demonstrations Dramatizations Role Playing Debates Panel Discussions Musical Recitals Movements
← Questioning →		
← Formal and Informal Observations →		
← Teacher/Student Dialogues →		

A laboratory practical exam, in which students are required to demonstrate safety procedures, might be an example of a Constructed Response-Performance assessment. Here, students might demonstrate appropriate pippetting skills, the proper way to dilute acids, or the safe way to heat a test tube. During these tasks, students would hardly be selecting their responses from a list; rather, they would be actively involved in demonstrating the acquired skills. Other performances that might occur in science classes include:

- Oral presentations of scientific findings (e.g., students present results of experiments or talk to judges at the science fair about their projects or students utilize whiteboards and multimedia displays to explain findings.)
- Dramatization of an important "moment in science" (e.g., the discovery of penicillin, invention of the microscope.)
- Taking on a role to solve a problem (e.g., being the "town member" or the "mosquito control board member" in a problem related to mosquito infestation.)
- Debates in which students support different sides of a question (e.g., stem cell research.)
- Panel discussions (i.e., each student on the panel is an expert in a particular area. The class attempts to solve a problem by calling on the panel for answers.)
- Musical recitals (e.g., students create raps or songs to teach scientific concepts.)

- Movements (e.g., students create an "electron dance" to show quantum levels or students become codons and dance protein synthesis.)

There are three other important methods of assessment depicted on the chart in **FIGURE 3.1**. These are *Questioning, Formal and Informal Observations*, and *Teacher/Student Dialogues*.

QUESTIONING is perhaps the most common type of classroom assessment. Brown and Edmondson (1984) reported that as many as 100 questions per class hour may be asked in elementary or secondary classrooms. For example, as teachers explain concepts, they often ask comprehension questions. A teacher explaining photosynthesis might ask at the end of the lesson, "What name is given to the process plants used to make food?" Or, he may ask, "What are the products of photosynthesis?"

Such comprehension questions, asking for a simple recall of facts comprise 70–80% of all questions asked in classrooms (Borich, 2000). If teachers only use questions to check comprehension, they are under-utilizing this valuable assessment method. Questions, like other assessment methods, can stimulate student thinking and thereby enhance learning. Such questions are deemed *effective* questions by Borich (2000, p. 238), who states that any question that "evokes a response that actively engages a student in the learning process" is an effective question.

RESOURCE

"Questions Are the Answer" This article by John E. Penick, Linda W. Crow, and Ronald J. Bonnstetter can be found in *Science Teacher, Volume 63, Number 1* (January 1996) on pages 26–29.

APPLICATION

Consider the following excerpt with accompanying comprehension (recall) questions. Answer each question. Then reflect: If students give correct answers, can you be sure they truly understand the concepts?

It is very important that you learn about mexicathon. Mexicathon is a new form of diodine. It is factographed in Guanocartera. The Guanocarterrans marborate large amounts of pavoline and then montisorate it to bractorize mexicathon. Mexicathon may well be one of our most symbonized chethorens in the future because of our fastulator zenorith.

1. What is mexicathon?
2. Where is mexicathon factographed?
3. How is mexicathon bractorized?
4. Why is it important to know about mexicathon?

To stimulate student thinking, questions must go beyond the comprehension level. Various authors categorize such higher-order thinking questions under different labels. For example, Bloom (1956) used the labels *Application, Analysis, Synthesis and Evaluation*. In revisiting Bloom's taxonomy, Anderson,et al., (2001) use *Apply, Analyze, Evaluate*, and *Create*. **FIGURE 3.2** uses the newer categorization scheme and provides sample science questions at each level. Note how the higher levels not only check to see if students can recall information, but also ask them to process information, clarify or expand on concepts, make judgments, formulate generalizations, and/or solve problems. Such questions take students forward in their thinking about science concepts (and therefore enhance learning), while still providing the teacher with a window into students' mental processes (serving an assessment function).

FIGURE 3.2
Higher-Order Thinking Questions for Eighth-Grade General Science

CATEGORY/ COGNITIVE LEVEL	SAMPLE SCIENCE QUESTIONS
REMEMBER	1. Which planet is closest to the sun? 2. Which planet is the largest one?
UNDERSTAND	1 Explain the Big Bang theory. 2. Explain the difference between weight and mass.
APPLY	1. The table on page 42 shows the distance from Earth to nine different stars. We know that light travels at a speed of 300,000 km per second. Would it take hours, days, or years for light from the Sun to reach Earth? From Alpha Centauri? 2. Use your definition of gravity and of orbital velocity to answer this question: We are launching two satellites that will orbit Earth. One will orbit at a much faster rate than the other. Which one (the slower or the faster satellite) will we be able to place further from the surface of Earth?
ANALYZE	1. Differentiate between meteors, comets, and asteroids, using at least three criteria. 2. In the excerpt we read, why does the astronaut say, "Then, on the Space Shuttle, a goat could be our best friend?"
EVALUATE	1. Of all the by-products of space exploration listed on page 93, which do you think is the most important? Why? 2. Do you agree or disagree with NASA's decision to ground the Space Shuttle? Defend your answer.
CREATE	1. In the story we read, astronauts decided to use a goat to replace the waste treatment plant on the Space Shuttle. What other animals could be used to replace heavy or bulky equipment on the Shuttle? 2. Design living quarters for one member of the crew on the Space Station.

FIGURE 3.2 emphasizes the need for teachers to create questions ahead of time—before instruction. Higher-order questions can be difficult to create, so teachers need time to plan these questions before they can implement the use of such higher-order questions in classes.

APPLICATION

Reflect on your own questioning skills. Thinking back on your last lesson, how would you categorize the questions you asked? How many went beyond the comprehension (*Remember, Understand*) category? For your next lesson, write out two questions for each cognitive level.

In addition to questioning techniques, teachers also use **OBSERVATIONS** as assessment methods. Sometimes these observations are formal ones; at other times they are informal. In *formal* observations, the student is usually notified ahead of time and given an opportunity to prepare for the observation. In addition, a set of criteria may be shared with the student to help him in this preparation. One example of the use of a formal observation might be during an oral presentation. The students would have time to prepare and perhaps be given a list of grading criteria as "Content," "Use of Visual Aids," "Organization," and "Clarity." The teacher then observes the student's performance and rates him on these criteria.

FIGURE 3.3 contains a sample data collection matrix for participation in seventh-grade class discussions of stem cell research. Following formal observations of students during such discussions, the teacher may record anecdotal notes on student performance. The teacher's findings can then be shared with the individual students to help them improve their participation in such classroom discussions. From the headings, it is obvious that the teacher is trying to emphasize the use of critical thinking in class discussions. She may be trying to enhance the cognitive challenge of such discussions by asking students to extend their knowledge beyond a recitation of facts. She wants them to interject inferences, justify their statements by referencing sources, apply knowledge, etc.

Data Collection Matrix—7th Grade

FIGURE 3.3 Participation in Class Discussions on Stem Cell Research

STUDENT'S NAME	CONTRIBUTED TO DISCUSSION	LISTENED TO OTHERS	ARTICULATED TEXT CONCEPTS IN OWN WORDS	SUBSTANTIATED IDEAS/ OPINIONS WITH REASONING	MADE INFERENCES	NOTES OR OBSERVATIONS

Adapted from Tools for Providing Feedback in Reading: A Reading Assessment Handbook for All Teachers in Grades 3–12, 2004, SERVE.

Other Sample Headings for Discussion Matrix:

- Differentiates important from extraneous information
- Relates text information to real life
- Applies textual facts to a new situation
- Makes accurate analogies

Informal observations occur more frequently than formal ones. Teachers are constantly watching their students to monitor understanding. A teacher who notes the frown on Martha's face during a lecture may stop and re-explain a concept. Circulating during seatwork, a teacher may find that Doug is not writing units after his numbers. She pauses to help him. Both of these teachers are using informal observations as assessment methods. They are using observations of students to help them ascertain what students know and are able to do as well as using observations to monitor the effectiveness of their own instruction.

In **FIGURE 3.3**, we saw how a formal observation instrument could be used to collect assessment data on student performance related to class discussions. Such class discussions would also fall under the category of Teacher/Student Dialogues. Let's look at another example of student/teacher interchanges in the following Classroom Scenario.

Classroom Scenario

In Ms. McCoy's tenth-grade chemistry class, students are classifying elements as solids, liquids, or gases at room temperature. It is Mary's turn, and Mary has been assigned the element mercury, which is a liquid at room temperature.

Ms. McCoy: OK, Mary, you have mercury. Is mercury a solid, liquid, or gas at room temperature?

Mary: Mercury is a solid.

(Teacher pauses and wonders if Mary has observed a broken mercury thermometer. When the thermometer breaks, the mercury liquid rolls up into small spheres that appear solid.)

Ms. McCoy: Please explain the reasoning for your answer, Mary.

Mary (*answering with some exasperation*): Mercury. Planet closest to the sun. Solid!

In this scenario, we see an example of teacher/student dialogue used as an assessment tool. Ms. McCoy wants to understand Mary's answer. She is trying to diagnose the reason Mary gave an "incorrect" answer. Ms. McCoy could simply have said, "Incorrect" and then given the right answer, "Mercury is a liquid at room temperature." However, she paused and asked for Mary's rationale. In doing so, she uncovered an important misconception that Mary held. Mary was equating Mercury the planet with mercury the element. Without this conversational exchange, Ms. McCoy would not know the reason for Mary's answer. However, now that she understands Mary's answer, she can help correct Mary's misconception. The use of teacher/student dialogues, then—probing for student understanding—can be a powerful assessment method. Other venues in which teacher/student dialogues occur include seminars, interviews, and individual conversations.

Aligning Assessment Methods with Learning Targets

In this chapter, we have explored several different assessment methods, including selected and constructed responses, questioning, observations, and teacher/student dialogues. Such assessments are only valid when they actually measure what they were intended to measure. To foster validity, it is important to tie assessments to particular learning targets. Let's select two sample benchmarks and then see if we design valid assessments for these targets. The source for both benchmarks is the South Carolina Science Curriculum Standards (**www.myscschools.com/offices/cso/standards/science**). To begin, we'll have to do some unpacking:

> **SAMPLE BENCHMARK 1:** Name major body parts and identify the uses of body parts.
> **KEY VERBS:** Name, identify
> **EMBEDDED LEARNING TARGET:** Knowledge

This sample benchmark has knowledge as its goal. Students must learn the names of the body parts and be able to recall their uses. For example, since this is a kindergarten-level benchmark, the students might be expected to learn that noses are for smelling, ears for hearing, mouths for eating, etc. What would be a good way to assess this benchmark? We know we have several options: selected response questions, constructed responses (either products or performances), teacher questioning, and teacher observation. **FIGURE 3.4** records some advantages and disadvantages of each assessment method.

FIGURE 3.4
Body Parts Assessment

ASSESSMENT METHOD	HOW THIS MIGHT LOOK	ADVANTAGES	DISADVANTAGES
SELECTED RESPONSE	Multiple-choice questions could refer students to a diagram of body parts, while matching questions could ask students to match body parts to uses.	Selected response questions can efficiently measure knowledge targets.	Due to limited reading/writing skills of kindergarten students, this is not a good match in this case. Also, as with all selected response questions, there is the possibility of false positive results due to chance.
CONSTRUCTED RESPONSE—PRODUCT	Students could be asked to draw and correctly label a diagram to assess knowledge of names of body parts. They could write uses beside each body part name.	An efficient method since teacher interactions with individuals is not required.	Again, due to limited reading/writing skills of kindergarten students, this is not a good match in this case.

ASSESSMENT METHOD	HOW THIS MIGHT LOOK	ADVANTAGES	DISADVANTAGES
CONSTRUCTED RESPONSE—PERFORMANCE	Students could don an apron on which body parts could be attached with Velcro (as stomach, lungs, heart, etc.). Students could correctly place organs on the apron, tell their names, and tell their uses.	Task is even more cognitively challenging than benchmark, as students must also correctly place the organ.	Students would need time to practice with the apron before the assessment. Time-consuming assessment, as each student will have to perform for the teacher.
TEACHER QUESTIONING COMBINED WITH TEACHER/STUDENT DIALOGUE	Teacher could point to particular body parts on a diagram and ask students to name them. A follow-up question could ask students to tell the use of this part.	Does not require reading or writing or even drawing. Should be within developmental level of students.	Time consuming. Again, the teacher would need to interact with every student to assess knowledge.
TEACHER OBSERVATION	This would be similar to the Questioning assessment. The student would point to a diagram, name the parts, and give the use of each. However, the teacher would simply watch the student, without using prompting questions.	Well within the development level of the students.	Prompting questions are not present, so the student may forget to give the name or give the use of the body part. Time consuming, requiring interaction between the teacher and each student.

As the chart shows, almost any assessment method would work for this benchmark. However, teachers must also take into consideration the developmental level of their students. Students with limited reading or writing ability may not handle certain types of assessments as well as others, particularly selected response or constructed response-product types of assessment. (Of course, there is no rule that says selected response questions must be written. The teacher could adapt this method and give oral choices from which the students may pick their answer. As long as the answer was given orally, this should still be well within the developmental level of kindergarteners.) Another way this could be tested is to have the teacher say the name of the body part and then have students circle the diagram showing this part. Since all the students could do this at their desks at the same time, this would not be as time consuming as performances, questioning, or observation methods. There is still the problem of giving the uses, however.

According to the chart in **FIGURE 3.4**, the best assessment methods for this benchmark were either Constructed Response-Performance or Teacher Questioning combined with Teacher/Student Dialogue. Both are time consuming, but both actually measure the knowledge embedded in the benchmark. The Constructed Response-Performance goes even further by stretching this minimum competency.

Let's look at another benchmark, first unpacking it and then looking at assessment choices:

SAMPLE BENCHMARK 2: Create and classify mixtures made of two or more substances (solid-solid, solid-liquid, and liquid-liquid).
KEY VERBS: create, classify
EMBEDDED TARGETS: Knowledge, Reasoning, Performance Skill

Within this fifth-grade benchmark, we find at least three learning targets embedded. Students must have some basic knowledge, as they must know the definitions of mixture, solid, and liquid. They will have to use reasoning to classify the types of mixtures. Finally, they must use performance skills to actually create a sample mixture of each type. So, what type of assessment should we use? Choose the best (most valid) one from the list below:

1) **Selected Response—Matching Question similar to this one.**

A. Solid-Solid	1. Ice water
B. Solid-Liquid	2. Brass
C. Liquid-Liquid	3. Oil and vinegar salad dressing
	4. Amalgam filling for a tooth

2) **Constructed Response—Product Short-Answer Question similar to this one.**

   Give one example of each type of mixture:

   Solid-Solid
   Solid-Liquid
   Liquid-Liquid

3) **Constructed Response—Performance Task with Teacher Observation similar to this one.**

   Using the materials in the lab, create three different mixtures. One must be Solid-Solid, one must be Solid-Liquid, and one must be Liquid-Liquid. Label your mixtures appropriately. Be ready to explain why these are mixtures, not compounds.

Hopefully, you would choose # 3 above as the most valid assessment. The other two do not assess the Performance Skill learning target, even though it was embedded in the benchmark.

APPLICATION

Choose a benchmark from your own state science standards. "Unpack" this benchmark into its component learning targets. Then, design a valid assessment for this benchmark.

In this chapter, we have concentrated on the various methods of assessments teachers may use to ascertain what students know and are able to do. We have particularly emphasized creating valid assessments—assessments that are aligned to the learning targets. Because the standards and benchmarks for a course may contain many different learning targets, a variety of assessments are needed. If teachers use only paper-and-pencil tests with selected response items, a rich source of data about student performance is missing.

In the next chapter, we will continue looking at a variety of assessments. The emphasis in Chapter Four shifts from defining assessment methods to providing meaningful feedback to students.

Gathering Evidence in a Variety of Ways

MEANINGFUL FEEDBACK

In Chapter Three, we learned that assessment methods can be categorised as *selected response* or *constructed response*. Constructed responses may be further divided into products and performances. Chapter Three also introduced Questioning, Observations, and Teacher/Student Dialogues as methods of assessment.

Gather Evidence in a Variety of Ways

In Chapter One, we stated that assessment has two broad goals: 1) to provide feedback to students about their own learning and 2) to monitor the effectiveness of instruction. We will tackle monitoring the effectiveness of instruction in Chapter Five. In this chapter, we will concentrate on providing feedback to students about their academic performance, using the above methods of assessment. We will primarily concentrate on assessments other than selected response types. The rationale for this arises from our work with teachers over the last five years. It appears that teachers already feel very comfortable utilizing selected response items. Teachers often have less experience, however, with other methods of assessment.

Purposes of Assessment

Reflect on your own experiences as a student. Was there a time when you received a low grade on an assignment because you weren't ready for an assessment? Perhaps you had not had sufficient time to practice the skill that was being assessed. Or, perhaps you THOUGHT you understood a concept, only to find on the test that you were mistaken. Assessment experiences like these happen to all of us. However, as good assessors, we try to limit the number of bad assessment experiences our own students must endure. To do this, we must first examine the three purposes of assessment: *diagnostic*, *formative*, and *summative*.

Teachers use **DIAGNOSTIC ASSESSMENTS** to ascertain the current level of understanding (or misunderstanding) that students possess about a concept prior to instruction on this concept. The chemistry teacher who gives a math test on exponents before beginning a unit on Avogadro's number (6.02×10^{23}) is using diagnostic assessment. This chemistry teacher wants to know if students can manipulate (add, subtract, multiply, and divide) exponents before beginning a unit that will require them to do so. Diagnostic assessments reveal "where the students are now" so that teachers can plan appropriate pathways that will lead the students to deeper levels of understanding. Such assessments are also useful in uncovering student misconceptions about science topics. Therefore, diagnostic assessments are powerful assessment tools that help teachers design instruction that meets students' needs.

FORMATIVE ASSESSMENTS are used to provide feedback to students as they progress toward a learning goal. Through formative assessments, students learn about their strengths and weaknesses before "it counts," i.e., before a grade is taken or recorded. Feedback during practice is the heart of formative assessment. For example, a sixth-grade student may submit the first draft of a graph for the teacher's review. The teacher makes comments on it or discusses the positive and negative features of the graph with the student. The student is then given an opportunity to re-draw or correct the graph before a grade is taken. Similarly, the teacher may provide a list of criteria to students. These criteria will be used to score a performance task (perhaps an oral presentation on "Animal Adaptations to a Particular Environment"). The students can practice the task, score themselves, and then ask peers to score them before performing the task for a grade.

In order for students to benefit from formative feedback, Black & Wiliam (1998, p.143) list the following three key features of effective feedback. "When anyone is trying to learn, feedback about the effort has three elements: recognition of the *desired goal*, evidence about *present position*, and some understanding of a *way to close the gap* between the two." Therefore, it is important for the teacher to convey his/her expectations to the students, provide evidence of their present level of performance, and then suggest strategies to improve that performance.

SUMMATIVE ASSESSMENTS are usually culminating ones. They occur after students have had time to practice a new skill, ask questions, interact with materials, etc. Summative assessments are generally used to rate the proficiency of students relative to a particular skill or skill set (as "proficient and ready to move on" to "beginning to understand, but needs further instruction") or they may be used to rank order students (as from "best" to "worst" or from "got it" to "hasn't a clue!").

In this chapter, we will focus on formative assessments. We will look at ways to provide meaningful feedback to science students *while they are learning*. In this formative process, we aid students in becoming competent self-assessors (since grading criteria are shared with them ahead of time), and we help them improve their performances (because they receive meaningful and timely feedback on those performances).

Types of Formative Feedback

Formative feedback is that feedback provided to students while they are learning. The purposes of such feedback include: 1) conveying the desired learning goal or goals to students, 2) alerting students to misconceptions, 3) providing them with information on strengths and weaknesses of their performances, and 4) providing strategies students can use to address weaknesses. Figure 4.1 displays the assessment methods introduced in Chapter Three and then suggests some feedback mechanisms appropriate for each. For the remainder of this chapter, we will explore examples of these feedback mechanisms.

FIGURE 4.1
Feedback Mechanisms Matched to Assessment Methods

ASSESSMENT METHOD	FEEDBACK MECHANISM
SELECTED RESPONSE	Written feedback: marking incorrect answers with circle or X; checkmarks for correct answers; Verbal feedback (follow-up conferences/discussions with students to review performance on selected response items)
CONSTRUCTED RESPONSE—PRODUCTS	Verbal feedback (discussions/conferences with students); Written feedback via comments or scoring rubrics
CONSTRUCTED RESPONSE—PERFORMANCES	Verbal feedback (discussions/conferences with students); Written feedback via comments or scoring rubrics
QUESTIONING	Verbal feedback (indicating correctness of answer; follow-up questions to stimulate thinking)
OBSERVATIONS—INFORMAL	Verbal feedback (via discussions/conferences with students; comments pinpointing weaknesses and strengths; suggestions for improving work)
OBSERVATIONS—FORMAL	Verbal feedback (via discussions/conferences with students); written feedback via comments, markings on observation instruments, or scoring rubrics
TEACHER/STUDENT DIALOGUES	Verbal feedback via comments highlighting strengths and weaknesses; suggestions for improving work

Verbal Feedback

Perhaps the most common method by which students receive feedback is via verbal comments from the teacher. This type of feedback is often used with Questioning, Teacher/Student Dialogues, and Informal Observations (although it is possible to use it with all kinds of assessments). A simple "Yes" or "That's correct" statement made by the teacher after a student answers a question may signal to the student that he is on the right track. Teacher comments after an observation, as "Trung, you need to hold the streaking wand lightly in your hand and use less pressure when streaking the agar plate. This will prevent cuts in the agar," can also provide students with feedback on performances as well as guidance for addressing weaknesses. To see how verbal feedback might be used in Teacher/Student Dialogues, let's return to the sample dialogue from Chapter Three. This time, we'll continue the conversation to see how Ms. McCoy provided verbal feedback to Mary.

Classroom Scenario

In Ms. McCoy's tenth-grade chemistry class, students are classifying elements as solids, liquids, or gases at room temperature. It is Mary's turn, and Mary has been assigned the element mercury, which is a liquid at room temperature.

Ms. McCoy: OK, Mary, you have mercury. Is mercury a solid, liquid, or gas at room temperature?

Mary: Mercury is a solid.

(*Teacher pauses and wonders if Mary has observed a broken mercury thermometer. When the thermometer breaks, the mercury liquid rolls up into small spheres that* appear *solid.*)

Ms. McCoy: Please explain the reasoning for your answer, Mary.

Mary (*answering with some exasperation*)**:** Mercury. Planet closest to the sun. Solid!

Verbal Feedback

Ms. McCoy: Thanks, Mary. Now I understand your answer. You're right—Mercury is the planet closest to the sun. However, the planet Mercury is not made up of the element "mercury." I know this is confusing, since both have the same name. So, let me give you and the class a little history lesson. Both the planet and the element are named after one of the Greek gods. This mythological god (whose name was Mercury) served as the messenger for the other gods. They would send him down to Earth to talk with mortals, or send him with messages to other gods. The reason Mercury was chosen to be the messenger was because he possessed sandals with attached wings. This made Mercury very fast. So, when the gods wanted to send a message, they naturally chose the fastest god as their messenger. When Mercury, the planet, was discovered by astronomers, they were struck by how fast the planet moved around the sun. Therefore, they chose to name this planet after the messenger of the gods. Mercury, the element, also got its name from being fast. In fact, the old name for this element was "quick-silver." Obviously, this element could move very fast at room temperature. Do you now understand that, even though both are named for the same mythological god, the planet Mercury and the element mercury are two different things?

Mary: Yes. I just thought that the planet had to be made up of the element, since both had the same name. Now, I see that they are just named after the same person, but they are two different things.

Ms. McCoy: Good, Mary. Are you ready to answer the question, then, about mercury, the element? Would it be a solid, liquid, or gas at room temperature?

Mary: You know, I'm still stumped. I just don't know which answer to choose.

Ms. McCoy: Then, let's review what we know about the Periodic Table of the Elements. How does the table in our book show us the different states of matter?

Mary: Oh, I remember! Solids are printed in black ink, liquids are in blue ink, and gases are in red ink.

Ms. McCoy: Absolutely correct, Mary. So, at room temperature, what is the state of matter of the element mercury?

Mary: Liquid!

Ms. McCoy: That's right, Mary. Mercury is a liquid at room temperature. We know this because our chart shows it in blue ink. If you look closely, you'll see that it is the only metal that is a liquid at room temperature.

In this interchange, Ms. McCoy provides feedback about Mary's answers. Initially, Mary gives the wrong answer, due to a misconception. Mary is confusing Mercury the planet with mercury the element. By expanding the Teacher/Student dialogue, Ms. McCoy gives Mary (and the class) some background knowledge about the element. She also carefully helps Mary get back on track with some skillful reminders and questions (e.g., How does the table in our book show us the different states of matter?). Once Mary is helped to the correct answer, Ms. McCoy provides positive feedback (e.g., "Absolutely correct, That's right") and then restates the right answer for Mary and the class (e.g., "Mercury is a liquid at room temperature"). Finally, Ms. McCoy provides some additional information that may help the class particularly remember the element mercury (e.g., "it is the only metal that is a liquid at room temperature").

In this classroom scenario, Ms. McCoy fulfills all of the Black and Wiliam (1998) requirements for feedback. She introduces the *desired goal* (students should be able to interpret the Periodic Table to ascertain which elements are liquids, solids, or gases at room temperature), and she provides information about Mary's *present position* relative to this goal through her verbal feedback. In the above interchange with Mary, we also see how Ms. McCoy reminds Mary of the strategy (*way to close the gap*) to use in formulating her answer (using the legend on the Periodic Table). Finally, Ms. McCoy also helps uncover a student misconception (the planet Mercury is made of mercury).

Written Feedback

Often, teachers use written feedback mechanisms when selected response and constructed response (product or performance) assessments are implemented with students. On selected response items, this written feedback is most likely to take the form of marking. As the chart in **FIGURE 4.1** indicates, these markings may take the form of circles, X marks, or checkmarks. These symbols provide a shorthand method of indicating correct or incorrect answers. Such marks do little, however, to help students close the gap.

Written feedback on products may also utilize marking symbols, but words or phrases are usually added to the paper to indicate student strengths or weaknesses or to provide guidance to students for improving performance. **FIGURE 4.2** displays an example of written feedback on an eighth-grade product assessment. In this case, the product is a short-answer, recall question given to students before an exam. The teacher is preparing the students for several knowledge-level questions that will be asked on a multiple-choice, large-scale test. By asking the students to provide a complete explanation of blood circulation, she can pinpoint weaknesses in their understanding. In this manner, students can correct misconceptions or identify areas for further study before the exam.

Written Feedback to a Short-Answer Question

FIGURE 4.2

QUESTION: Trace the path of a single red blood cell in the human circulatory system. Start at the aorta. Note when the blood is oxygen-rich and when it is oxygen poor.

STUDENT ANSWER (with teacher feedback in green)**:**

Oxygen rich blood leaves the heart through the aorta. It goes to capillaries where it gives up some of its oxygen to tissues. From the capillaries, the blood flows into veins which take the blood back to the heart. The heart pumps the blood to the lungs, where it picks up more oxygen and then returns to the heart. The heart then pumps the blood back into the aorta for another round.

Teacher feedback:

- Correct! The blood is oxygen rich here.
- Gives up oxygen to tissues where?
- Between the aorta and the capillaries. Where does the blood go?
- Name of the major vein that returns blood to the heart?
- Which part of the heart?
- Correct
- Where in the lungs does this occur?
- Through which type of vessel?

Keisha, You've made a good start here. You need to add more details, particularly about the parts of the heart and the types of blood vessels. Remember our saying, "Artery starts with "A" and carries Blood Away." When will this saying not be true? (There is one time!)

In the written feedback, the teacher indicates the strengths and weaknesses found within the student's answer. She appears to use "Correct" to provide positive feedback and questions to indicate weaknesses. She ends her comments with suggestions about how Keisha can improve her performance on this subject (blood circulation) and even challenges Keisha to perform further research/study ("When will this saying ***not*** be true?").

In this example of written feedback, then, the teacher identifies the desired goal (students will be able to trace the course of a red blood cell throughout the circulatory system), provides information on student strengths and weaknesses, and suggests strategies for addressing weaknesses. This particular excerpt does not indicate any misconceptions found in Keisha's thinking. It appears that Keisha has a basic understanding of the circulatory system but needs to work on the details. Knowing the names of particular blood vessels (as the vena cavae), the parts of the heart (right and left atria, right and left ventricles), and the functions of each component part of the circulatory system will prepare Keisha for a multitude of highly specific, knowledge-level, multiple-choice questions on this subject. In this manner, the short-answer question is formative in nature, preparing Keisha for a summative assessment that will come later. The feedback is formative as it occurs while Keisha is still learning and because it fulfills three of the four purposes of such feedback.

A Subset of Written Feedback: Scoring Rubrics

Scoring rubrics are a special subset of written feedback. Rubrics generally provide a list of criteria supported by descriptions of performance levels relative to these criteria. **FIGURE 4.3** provides an excerpt from such a scoring rubric, showing one criterion with its accompanying performance descriptions. This rubric might be used in a high school biology class. In this class, the teacher uses a machine to simulate the heart sounds that would be heard through a stethoscope during a blood pressure reading. Students must identify when to take readings by listening to the heart sounds.

FIGURE 4.3
Excerpt from Rubric Used to Assess Student Skill in Taking Blood Pressure

CRITERIA	EXCELLENT	ADEQUATE	NEEDS IMPROVEMENT
ACCURACY	Student readings for systolic and diastolic pressure are within 3 mmHg (+ or -) of machine setting	Student readings for systolic and diastolic pressure are within 7 mmHg (+ or -) of machine setting	Student readings for systolic and diastolic pressure are greater than 7 mmHg (+ or -) compared to the machine setting

This rubric provides feedback to students by stating the desired goal (an acceptable performance is to be within 7 mmHg of the machine reading) and by providing information about student performances (students can score at the "excellent," "adequate," or "needs improvement" level). This particular excerpt from the rubric does not provide information on strategies students can use to close the gap between "excellent" and "needs improvement" performances, however.

There are distinct advantages to using rubrics for formative feedback. When rubrics are distributed to students while they are learning, the rubrics can be very instructional in nature. Because such rubrics list the criteria by which students will be judged and then provide descriptions of performances at different levels, students can use them to hone their own performances. Therefore, rubrics inform students of the desired goals and provide mechanisms by which students can self-assess (gain information on their present position). Rubrics help students plan their performances, as they provide clues as to what a quality performance will include. These clues can often be used to brainstorm strategies for improving performance.

By using a rubric, the teacher provides the students with her expectations. She defines and describes a high-quality performance and lists the criteria by which the students will be judged. She distributes the rubric before students begin work. Then, she asks the students to perform peer and self-assessments of the work, using the rubric as a guide. Finally, she provides feedback on the student work by scoring this with the rubric. All of these activities (peer, self, and teacher assessment) can be formative—they can take place during the learning process. Therefore, the rubric serves as a feedback tool during this formative process. However, once students have completed the peer, self, and teacher formative assessments, the same rubric can then be used to score a summative assessment. In this manner, the same rubric may be used both formatively and summatively. In this chapter, we will emphasize the formative use of rubrics, however. We are drawing a distinction between "scoring" an assignment and "grading" an assignment. For the purposes of this text, "scoring" will mean giving formative feedback for improvement during instruction, whereas "grading" connotes a summative function that occurs after instruction.

Whether used for formative purposes or summative ones, rubrics may be *holistic* or *analytical* in nature. It is usually very easy to distinguish these two types. Students will receive only one score when a holistic rubric is used, but will receive multiple scores within several different dimensions when an analytical rubric is used. See **FIGURE 4.4** for sample holistic and analytical rubrics.

FIGURE 4.4
Types of Rubrics

Holistic Rubric for Seventh-Grade Writing in Science

SCORE POINTS AND LEVEL	DESCRIPTION
6: HIGHEST LEVEL	Main ideas and supporting details are entirely accurate, specific, clear, and sufficient; Higher-order thinking highly developed; Strong organizational structures throughout; Strong use of language throughout; Diagrams or sketches are accurate, clear, and focused on the topic
4: SOMEWHAT DEVELOPED RESPONSES	Main ideas and supporting details are mostly accurate and mostly specific; Higher-order thinking developed; Satisfactory organizational structures; Language used in a satisfactory manner; Diagrams or sketches are on the topic
2: UNDEVELOPED RESPONSE	Main idea with no support and/or contains major errors in using the main science concepts important to the task; Very little evidence of higher-order thinking; Unorganized; Some language used incorrectly; Diagrams or sketches are off topic or absent
0: LOWEST LEVEL	No response

Analytical Rubric for Seventh-Grade Writing in Science

CRITERIA	LEVEL 1—LOW	LEVEL 2	LEVEL 3	LEVEL 4—HIGH
MAIN IDEAS	Not present	Present but not clear; May include major errors	Clear and accurate; May include minor errors	Clear, accurate, and sufficient in number to cover the topic
SUPPORTING DETAILS	Not accurate	Accurate but very sparse; May include major errors	Accurate and sufficient in number to support the main ideas; May include minor errors	Accurate, sufficient in number, and especially well chosen to support the main idea
HIGHER-ORDER THINKING	Not evident	Somewhat evident	Evident	Strongly evident
ORGANIZATION	Very unorganized	Somewhat organized	Organized	Strongly organized throughout
USE OF LANGUAGE	Immature use of vocabulary	Vocabulary includes some topic-specific words	Vocabulary includes a good selection of topic-specific words	Vocabulary chosen does an excellent job of addressing both the topic and the audience
SKETCHES	No sketches	Sketches present but do not support the writing well	Sketches accurate and do support the writing	Sketches accurate and especially clear in supporting the writing

Adapted from: Hibbard, K. M. (2000). Performance-based learning and assessment in middle school science. Larchmont, NY: Eye on Education. *Pages 70–71, 74.*

From examining these two rubrics, you can see that the student would receive only one score in the first, holistic rubric. She would get either 6 or 4 or 2 or 0. However, in the second, analytical rubric, the student would get a score in six different dimensions (Main Idea, Supporting Details, Higher-Order Thinking, Organization, Use of Language, and Sketches).

Let us suppose that a seventh-grade student, Marty, submitted a sample of scientific writing. In this piece, Marty covered the main idea very well, but was weak in supporting details. He showed exceptional higher-order thinking, as he related the science concept to his own life and extrapolated possible future uses of this concept. However, his organization tended to skip from subject to subject and back again, negatively impacting his clarity. His word choices were high level. Marty forgot to include a sketch. How would Marty do on each of the rubrics?

Perhaps the teacher using the holistic rubric gave Marty an overall 4 due to the poor organization and lack of sketch. The teacher using the analytical rubric recorded the following for Marty:

Main Idea	4
Supporting Details	2
Higher-Order Thinking	4
Organization	3
Use of Language	4
Sketches	1

Which type of rubric would provide the most meaningful formative feedback to Marty, to help him improve his scientific writing performance? Certainly it would be the analytical one, which clearly shows Marty's strengths and weaknesses. This analytical rubric could greatly aid Marty in self-assessing his own work by informing him of targets, indicating his present position relative to the targets, and suggesting ways to improve his performance (via performance indicators on the rubric, as comparing the descriptions of performance levels for Supporting Details). By reading the rubric, Marty could find that he needs to provide more supporting details (his were sparse) and that these details must support his main idea. The rubric also notifies Marty of the necessity of enclosing a sketch with his work. Marty can use this information to improve his performance before a grade is taken on his writing in science. Since the analytical rubric provides greater details on student performance (giving a score in several different dimensions to help students pinpoint strengths and weaknesses), we recommend the use of *analytical* rubrics with formative assessments.

Creating High-Quality Analytical Rubrics

Anyone who has attempted to create a rubric from "scratch" has probably found that this is no easy task. It is not unusual for teachers creating rubrics to overlook an important criterion and only realize they did so when they use the rubric to score student work. For example, one teacher created a rubric for assessing student posters on types of muscles. Her criteria included: Clarity, Use of Color, Size, and Depiction of Three Types. Only when she used the rubric to score student posters did she realize she had not included Accuracy as a criterion. Therefore, if students labeled the types of muscle incorrectly (perhaps labeling cardiac muscle as skeletal muscle), nowhere in her rubric could she count off for this! To avoid such embarrassing experiences, we recommend that you create new rubrics while examining student work from past years. One way to begin this process is to sort the work into piles labeled "Low, Medium, and High." Then, by looking at the differences in the work, you can begin to construct a list of important criteria. After you have the criteria, you can write descriptions for the different proficiency levels.

Another excellent way to create rubrics is to involve students in the development process. One reason for doing so is that many minds are better than one: Students may think of criteria that the teacher would miss. Another important reason for involving students in rubric creation is simply because of the educative value of this process. By discussing the criteria for quality work, students will get a clearer idea of the teacher's expectations. Whether rubrics are teacher generated or student constructed, it is imperative to keep the learning target to the forefront, so that the rubric accurately reflects what students are being asked to accomplish.

FIGURE 4.5 has a confusing name, but it may help you understand the important characteristics of a high-quality rubric. It is a rubric that can be used to judge the quality of an analytical rubric that you create for use in your classroom. If you score your teacher-constructed classroom rubric using **FIGURE 4.5**, you may find weaknesses in the classroom rubric that you can correct before you disseminate it to students. By using **FIGURE 4.5**, you may enhance the quality of the rubric.

FIGURE 4.5
A Rubric for Assessing the Quality of an Analytical Rubric

	CRITERIA		
	Excellent (5)	**Acceptable (3)**	**Unacceptable (1)**
1	Criteria identify all components necessary for high-quality work.	Sufficient criteria are present to define high-quality work, but some "fine-tuning" criteria may be missing.	Rubric fails to include critical criteria for excellence (eg., content accuracy).
2	Criteria reflect alignment to all stated objectives.	Minor objectives are not sufficiently addressed by the criteria.	Major objectives are not sufficiently addressed by the criteria.
3	Criteria are clearly feasible and measurable.	Criteria are measurable and reasonably feasible.	Criteria listed in the rubric may not be measurable, or they have low feasibility (e.g., they would require the teacher to individually observe each student, thereby using more time than warranted by the importance of the criterion).

FIGURE 4.5 (continued)

	DESCRIPTORS		
	Excellent (5)	**Acceptable (3)**	**Unacceptable (1)**
4	Descriptors always use "kid-friendly" language, language easily understood by the target student population.	Descriptors usually use "kid-friendly language," but some "educationalese" is present.	Descriptors consistently use language inappropriate for the developmental/academic/grade level of the student population.
5	Descriptors use observable characteristics and avoid overuse of vague quantitative terms (e.g., several, numerous, some).	Descriptors rely more on quantitative terms than on observable characteristics.	The use of too many generalities (as sometimes, occasionally, some, etc.) detracts from the clarity and leads to confusion.
6	Descriptors accurately and specifically describe levels of performance within the criterion.	Descriptors accurately describe levels of performance, but specificity may be low (e.g., for a writing rubric, the descriptors might say the student "uses standard English." This does not specifically explain what "standard English" means).	Descriptors and criteria do not match OR there are no descriptors OR descriptors use one word (e.g., "always" for best performance, "sometimes" for acceptable performance, and "never" for unacceptable performance).

	SCORING GUIDE		
	Excellent (5)	**Acceptable (3)**	**Unacceptable (1)**
7	The most important criteria are allotted the most number of points, thereby allowing students to prioritize efforts.	All criteria earn the same amount of points, as all are equally important.	All criteria earn the same amount of points, but some are clearly more important or more aligned to the objectives than others.
8	The scoring guide clearly demonstrates how rubric scores will be converted to grades.	A conversion scale is present, but may not address all grading levels (e.g. "10 points is an A" is stated, but no distinctions are given for A+ grades or A- grades).	No conversion scale is present.

	FORMAT (TYPE) OF RUBRIC		
	Excellent (5)	**Acceptable (3)**	**Unacceptable (1)**
9	Type of rubric used supports purpose of assessment (e.g., an analytical, task-specific rubric for a formative assessment; a holistic, generalized rubric for a summative assessment).	Type of rubric used is appropriate (e.g., analytical for formative assessments and holistic only for summative assessments), but generalized may be substituted for task-specific rubrics if students will perform several iterations of this type of work throughout the school year (e.g., an "oral presentation" rubric).	Type of rubric does not support purpose of assessment (e.g., a holistic rubric is used to judge performance on a newly taught skill).

QUALITY OF THE FEEDBACK			
	Excellent (5)	Acceptable (3)	Unacceptable (1)
10	Rubric scores clearly pinpoint student's strengths and weaknesses; student obtains sufficient information on strategies that can be used to improve performance.	Rubric scores clearly pinpoint student's strengths and weaknesses and provide some information on improvement strategies.	Rubric gives insufficient information to allow student to improve.
11	Rubric is instructional in nature; students can read and understand what high-quality work will look like; students can use rubric to self-assess own performance without input from the teacher.	Rubric is instructional in nature; students can read and understand what high-quality work will look like.	Rubric does not contain sufficient information for students to understand teacher expectations and/or what quality work will look like.

Sample Rubrics

APPLICATION

Select a rubric you use in your science classroom. Score this rubric using Figure 4.5. Redesign the rubric based upon your findings.

In this final section of Chapter Four, we have included sample rubrics for several tasks typically found in science classrooms. The rubrics we have chosen are ones that are analytical in nature, in hopes that they will be used to provide meaningful feedback to students during the learning process. The rubrics are designed for use in a high school classroom. Each rubric also includes a sample score conversion chart. It should be emphasized, however, that the purpose of providing such scores is informational, rather than evaluative in nature. The scores show the students the proficiency level at which they are currently working. They allow the students to predict a grade based on current performance, in hopes of encouraging improvement in this performance. Since the rubrics included here are *formative* in nature and because they are being used to provide feedback to students *during the learning process*, grades would not actually be recorded. These same rubrics could be used summatively, however, after students are given sufficient practice time and quality feedback on their practice attempts.

Cautionary notes: The following rubrics are samples only. They are not ready-made rubrics available for instant use in your classroom. Teachers should use caution in using these or any other published rubrics. Rubrics should always be carefully evaluated for alignment with the curriculum being used. Since different states have different standards, it is important for the readers of this manual to understand that the sample rubrics included here may provide a better fit for some standards than for others.

Each sample rubric is preceded by a brief explanation of the task students are being asked to perform. Connections to content standards from the National Science Education Standards are also provided for each rubric. Before incorporating one of these sample rubrics in your own classroom, you should first evaluate it using Figure 4.5. If the rubric does not align with the standards used in your classroom, it will need to be revised before you utilize it. If the greatest number of points are awarded to criteria that are of little value in your curriculum, the point values should also be adjusted. Finally, the sample rubrics all use a grade of "C" to indicate mastery of the task. If you wish to set a higher grade for mastery, you may need to indicate this in the score conversion chart by having students resubmit work below your mastery level.

The sample rubrics, like all rubrics, are not meant to be used as stand-alone instructional tools for students. While these rubrics are instructional in that they provide lists of criteria and descriptions of performance levels, students will need support in understanding the descriptions and in utilizing the rubrics. When introducing a new rubric to the class (particularly one that was primarily teacher-generated without input from students), the following suggestions may be helpful:

1) Set aside time within class for students to read through the rubric.
2) Highlight the learning target to be scored with this rubric.
3) Review each criterion, with its accompanying performance level descriptions with the students. During this review, ask students to put the descriptions in their own words and ask them to define unfamiliar vocabulary. If they are unable to perform these tasks, more teacher explanations about the rubric are needed.
4) Provide a sample (or samples) of student work that students can score by using the rubric. Have students share their scores and explain why particular scores were given, using "rubric language" (language from the rubric) to justify the scores.
5) Work together with students to revise the rubric. Change language to be more kid-friendly, based upon student recommendations. Add or delete criteria as needed to more closely align the rubric with the learning target.
6) When using the rubric to score formative assignments, don't hesitate to add written comments. Don't feel constrained to simply circle particular scores for each dimension or to highlight rubric phrases. Teacher comments may help students understand strategies they can use to improve performance.

APPLICATION

Use the "Rubric for Assessing the Quality of an Analytical Rubric" found in Figure 4.5 to evaluate one or more of the sample rubrics found in Figures 4.6–4.9. What are the strengths and weaknesses of these rubrics?

With the above cautions in mind, review the sample rubrics with a critical eye. Use the following Application suggestions to enhance the utility of each rubric in your science class.

APPLICATION

For each rubric:

1. Rewrite, using language that students at the level you teach would understand ("kid language") instead of the high school language incorporated in these rubrics.
2. Revise the rubric as needed to align with your curriculum.
3. Check the rubric to ensure that the score conversion chart accurately portrays mastery levels. (For example, in the laboratory report rubric, Figure 4.9, would 34 points—the "C" level—correspond to mastery? If not, revise the score converstion chart.)
4. Test the rubric by using it with your students.
5. Rewrite the rubric to address any shortcomings you found upon implementation or to make the rubric more relevant to your classroom.
6. Check the final rubric to ensure that it fulfills the purposes of formative feedback which include: a) conveying the desired learning goal or goals to students, b) alerting students to misconceptions, c) providing them with information on strengths and weaknesses of their performances, and d) providing strategies students can use to address weaknesses.

Designing an Experiment

In science, students are often asked to design an experiment in order to answer a question or to solve a problem. The rubric displayed in Figure 4.6 may be used to provide feedback on this design process. This rubric might be used when students are given "mystery powders" as well as several testing solutions and then asked to create a process for identifying each powder. With the modification suggested at the bottom, this rubric could also be adapted to provide feedback on such engineering design tasks as creating a "cradle" that will protect the egg from breaking in the classic "egg drop" experiment.

RESOURCES

Websites for Rubrics:

- www.accessexcellence.org
- http://rubistar.4teachers.org/index.php
- www.rubrics4teachers.com
- http://intranet.cps.k12.il.us/Assessments/Ideas_and_Rubrics/Rubric_Bank/rubric_bank.html

This rubric may be used to assess the following in grades 9–12 (National Research Council, 1996, p. 175):

- Identify questions and concepts that guide scientific investigations.
- Design and conduct scientific investigations.

FIGURE 4.6 — **Generic Experimental Design Rubric**

CRITERIA	EXEMPLARY	ACCEPTABLE	RESUBMIT
SCIENCE KNOWLEDGE	**10 points** Design accurately and completely summarizes scientific concepts/facts/theories relevant to the experiment. Design explicitly makes connections between this scientific information and the experimental design.	**7 points** Design includes essential scientific concepts/facts/theories but may not provide complete coverage of topic. While some connections between scientific knowledge and experimental design are explicit, insufficient explanations are present for the reader to completely understand the basis for choosing the experimental design.	**4 points** Design includes some scientific concepts/facts/theories, but some essential items are missing (insufficient information for reader to gain a basic understanding of the topic). Connections between this information and experimental design are missing.
PROBLEM	**7 points** Problem is clearly stated as a question, is measurable, and it is feasible that a solution could be reached on the problem within the constraints of available equipment, environment, and time. Predictions/hypotheses relevant to the problem are included in the design.	**5 points** Problem is stated as a question and is measurable, but feasibility is low for solving the problem (it requires unusual equipment, special environments, long periods of time, etc.), as the student may fail to adequately explain how special conditions will be obtained. Predictions/hypotheses are relevant, but may not be completely explained.	**3 points** Problem is not measurable and is not feasible as student makes no attempt to explain how special, required conditions will be met. Predictions/hypotheses are absent or irrelevant.
CONTROLS AND VARIABLES	**10 points** Design clearly describes scientific controls and acknowledges/explains the necessity for the controls. Design identifies independent and dependent variables correctly.	**7 points** Design clearly describes scientific controls, and the design identifies independent and dependent variables correctly.	**4 points** Controls inadequately explained and/or missing *OR* design fails to identify independent and dependent variables correctly.

CRITERIA	EXEMPLARY	ACCEPTABLE	RESUBMIT
DATA COLLECTION PLAN	**5 points** Includes a materials/equipment list. Materials/equipment are appropriate to the task and are the best available items to use in order to collect data relevant to the problem.	**3 points** Includes a materials/equipment list. Materials/equipment are appropriate to the task, but may not be the best available items to use in order to collect relevant data.	**1 point** No materials/equipment list present *OR* list is incomplete *OR* materials/equipment listed are not suited to collect data relevant to stated problem.
PROCEDURE	**12 points** Steps in procedure are numbered and sequenced logically. Sufficient explicit details are included to allow the reader to duplicate the experiment. Illustrations (as for set-ups of equipment) are accurate and would aid the reader in duplicating the experiment. Steps outlined in the procedure align to stated problem and clearly have the potential to answer the experimental question. Safety rules are accurately incorporated into procedural steps. Sufficient trials or multiple tests are present to support (or not support) the predictions/hypotheses.	**9 points** Steps in procedure are sequenced logically, and enough information is present to duplicate the experiment. Steps align to the problem and have the potential to answer the question. Safety rules are accurately incorporated into procedural steps. Sufficient trials are present.	**5 points** Steps are not sequenced logically *OR* procedure does not provide enough details to enable duplication of the experiment *OR* safety rules are not incorporated into the procedure *OR* sufficient trials are not present.
CONCLUSION PLAN	**5 points** Design designates the manner in which findings will be graphically displayed (table, line graph, bar graph, pie chart, etc.). Graphic is best match for displaying this type of data.	**3 points** Design designates the manner in which findings will be graphically displayed. Graphic chosen is adequate for the task.	**1 point** Design fails to designate manner in which findings will be graphically displayed OR graphic chosen is inadequate for the task.

CRITERIA	EXEMPLARY	ACCEPTABLE	RESUBMIT
SELF-ASSESSMENT	**5 points** Student uses this rubric to assess the design and gives self a score in each dimension of the rubric. Student writes a short paragraph highlighting strengths and weaknesses of the experimental design. In the paragraph, the student provides accurate rationale, with supporting details (using language from the rubric), for the classification of the strengths and weaknesses.	**3 points** Student uses this rubric to assess the design and gives self a score in each dimension. Student writes a short paragraph highlighting strengths and weaknesses of the experimental design. Although not stated, the rationale for classifying strengths and weaknesses is clearly related to those described in the rubric.	**1 point** Student fails to give self scores in each dimension of the rubric *OR* fails to write a paragraph highlighting strengths and weaknesses *OR* the strengths and weaknesses described by the student do not relate to those described by the rubric.

Optional Additions for Engineering-type Experimental Design (e.g., “egg-drop” experiment)

DESIGN SPECIFICATIONS	**5 points** Procedure incorporates (and never violates) the outlined requirements or specifications for the designed product.	**5 points** Procedure incorporates (and never violates) the outlined requirements or specifications for the designed product.	**1 point** Procedure violates the outlined requirements or specifications for the designed product.

Score Conversion (without Engineering Design criterion)

48–54 → A   40–47 → B   30–39 → C   Below 30 → Resubmit

Presenting Scientific Findings

In science, students are often called upon to present their work to others. This is an important aspect of science, as scientists must share their findings with their colleagues in order to advance scientific knowledge. The rubric in **FIGURE 4.7** provides criteria for a multimedia slide presentation of scientific work, while **FIGURE 4.8** describes important oral presentation skills.

The multimedia slide presentation rubric may be used in assessing the following expectation for students in grades 9–12 (National Research Council, 1996, p. 175):

- Students will use technology to improve investigations and communications.

Both the multimedia slide presentation rubric and the oral presentation rubric could be used in assessing the following expectation:

- Students will communicate and defend a scientific argument.

FIGURE 4.7
Rubric for the Technology Slide Presentation in Science

Note: Oral presentation of slides will be assessed using the Oral Presentation Rubric

CRITERIA	POINTS POSSIBLE	POINTS AWARDED	EXEMPLARY	INTERMEDIATE	ACCEPTABLE	UNACCEPTABLE—RESUBMIT
CONTENT	**20**		**20 points** Scientific information on slides is accurate and covers topic completely and in depth. All essential facts (ones needed to foster understanding of the topic) are present.	**15–17 points** Scientific information on slides is generally accurate (no more than 2 errors in entire presentation), and most essential facts are present. Topic not completely covered, but includes enough elaboration that audience will gain a fair understanding of the topic.	**10–14 points** Includes little essential information and/or scientific facts presented may be inaccurate (2–4 errors in entire presentation). Incomplete coverage, but audience may still gain a basic understanding of the topic.	**0 points** Inadequate coverage and/or more than 4 factual errors and/or most essential facts missing.

CRITERIA	POINTS POSSIBLE	POINTS AWARDED	EXEMPLARY	INTERMEDIATE	ACCEPTABLE	UNACCEPTABLE—RESUBMIT
SUBSTANTIATION	**15**		**15 points** Science facts are substantiated with on-slide citations and with accurate references slides at the end of the presentation. All needed citations are present. (They are needed if the information is not a personal idea, but borrowed, quoted, or adapted from another source.) All citations are encompassed in the reference slides. Citations and references use APA style correctly.	**12–14 points** 1–2 needed citations are missing OR 1–2 citations left out of reference list OR 1–2 citations or references use incorrect APA style.	**9–11 points** 3–4 needed citations are missing *OR* 3–4 citations left out of reference list *OR* 3–4 citations or references use incorrect APA style.	**0 points** More than 4 needed citations are missing *OR* more than 4 citations left out of reference list *OR* more than 4 citations or references use incorrect APA style.
LAYOUT/DESIGN	**10**		**10 points** Titles use 44 size font; subtitles use 36 size. No information on slide is less than 26 size. Font style chosen enhances legibility; colors enhance readability, rather than interfere with it. Format is consistent throughout presentation (e.g., important terms in red; bullets all same format; etc.) Transitions and animations are purposeful. They contribute to an understanding of the science topic rather than slowing the presentation, interfering with understanding, or simply adding "bells and whistles" for no apparent reason.	**8–9 points** Correct font sizes used, font style is legible, original color enhances readability, but "followed links" colors may make words hard to discern. There may be 1–2 inconsistencies in format. Transitions and animations are generally purposeful, although 1–2 may fail to contribute to the understanding of the science topic.	**5–7 points** 1–4 errors in font size, OR colors used in original text interfere with readability, OR font style detracts rather than enhances legibility, OR 3–5 inconsistencies in format.	**0 points** More than 4 errors in font size, *OR* colors used in original text make words illegible, *OR* font style makes words illegible, OR more than 5 inconsistencies in format.

CRITERIA	POINTS POSSIBLE	POINTS AWARDED	EXEMPLARY	INTERMEDIATE	ACCEPTABLE	UNACCEPTABLE—RESUBMIT
ORGANIZATION	**10**		**10 points** Sequence chosen enhances the understanding of the science topic. This may be from specific to general, general to specific, chronological, etc. Science knowledge builds as audience moves through slides.	**8–9 points** 1–2 slides depart from overall organizational scheme, thereby detracting from knowledge construction by audience.	**5–7 points** 3–4 slides depart from overall organizational scheme, thereby detracting from knowledge construction by audience.	**0 points** No discernable organizational scheme.
GRAPHICS	**10**		**10 points** Each slide contains at least one graphic. Graphics on a particular slide help illuminate (help audience understand) the information on the slide. There is a mix of graphic types within the presentation (mix of photos, graphics, tables, clip art, icons, symbols, student-created, etc.). Graphic size is appropriate (not too big or too small) and resolution is good.	**8–9 points** There is a good mix of graphic types within the presentation. 1–2 slides contain no graphics OR 1–2 slides contain graphics that fail to illuminate the information on the slide OR 1–2 graphic sizes or resolutions are inappropriate.	**5–7 points** Slides primarily display only 4–5 types of graphics OR 3–5 contain no graphics OR 3–5 contain graphics that fail to illuminate the information on the slide OR 3—graphic sizes or resolutions are inappropriate.	**0 points** Slides display 3 or less types of graphics *OR* more than 5 slides contain no graphics *OR* more than 5 slides contain graphics that fail to illuminate the information on the slide *OR* more than 5 graphic sizes or resolutions are inappropriate.

CRITERIA	POINTS POSSIBLE	POINTS AWARDED	EXEMPLARY	INTERMEDIATE	ACCEPTABLE	UNACCEPTABLE—RESUBMIT
MECHANICS	**5**		**5 points** Grammar, spelling, punctuation, capitalization are all correct. No other mechanical errors (subject/verb agreement, incomplete sentences, etc.) are present.	**3–4 points** 1–2 mechanical errors	**1–2 points** 3–4 mechanical errors	**0 points** 5 or more mechanical errors

Score Conversion (without Engineering Design criterion)

55–70 → A   54–62 → B   30–48 → C   Below 30 → Redo

FIGURE 4.8
Oral Presentation in Science

PERFORMANCE ELEMENT	POINTS POSSIBLE	POINTS AWARDED	EXEMPLARY	GOOD	ADEQUATE	RESUBMIT
AWARENESS OF AUDIENCE	**10**		**10 points** Significantly increases audience understanding and knowledge of topic; Effectively convinces an audience to recognize the validity of a point of view.	**8 points** Raises audience understanding and awareness of most points; Clear point of view, but development or support is inconclusive and incomplete.	**6 points** Raises audience understanding and knowledge of some points; Point of view may be clear, but lacks development or support.	**4 points** Fails to increase audience understanding or knowledge of topic. Fails to effectively convince the audience.

PERFORMANCE ELEMENT	POINTS POSSIBLE	POINTS AWARDED	EXEMPLARY	GOOD	ADEQUATE	RESUBMIT
STRENGTH OF MATERIAL AND ORGANIZATION	**20**		**20 points** Clear purpose and subject; Pertinent examples, facts, and/or statistics; Conclusions/ideas are supported by evidence; Major ideas summarized and audience left with full understanding of presenter's position.	**15 points** Has some success defining purpose and subject; Some examples, facts, and/or statistics support the subject; Includes some data or evidence to support conclusions or ideas; May need to refine summary or final idea.	**10 points** Attempts to define purpose and subject; Weak examples, facts, and/or statistics, which do not adequately support the subject; Includes very thin data or evidence in support of ideas or conclusions; Major ideas may need to be summarized or audience is left with vague idea.	**5 points** Subject and purpose are not clearly defined; Very weak or no support of subject through use of examples, facts, and/or statistics; Totally insufficient support for ideas or conclusions. Major ideas left unclear, audience left with no new ideas.
DELIVERY	**20**		**20 points** Relaxed, self-confident and appropriately dressed for purpose or audience; Builds trust and holds attention by direct eye contact with all parts of audience; Fluctuations in volume and inflection help to maintain audience interest and emphasize key points;	**15 points** Quick recovery from minor mistakes; Appropriately dressed; Fairly consistent use of direct eye contact with audience; Satisfactory variation of volume and inflection.	**10 points** Some tension or indifference apparent and possible inappropriate dress for purpose or audience; Occasional but unsustained eye contact with audience; Uneven volume with little or no inflection.	**5 points** Nervous tension obvious and/or inappropriately dressed for purpose or audience: No effort to make eye contact with audience; Low volume and/or monotonous tone cause audience to disengage.

Score Conversion

44–50 → A   38–43 → B   26–37 → C   19–25 → D   Below 19 → Resubmit

Writing a Lab Report

Another common task given to science students is to write a laboratory report of an experiment. The rubric shown in Figure 4.9 would provide meaningful feedback to high school students writing such reports.

This rubric addresses parts of the following (National Research Council, 1996 pp. 174–176):

- Design and conduct scientific investigation.
- Use technology and mathematics to improve investigations and communications.
- Formulate and revise scientific explanations and models using logic and evidence.
- Communicate and defend a scientific argument.

RESOURCE

The following website may be useful for high school teachers. This "Lab Write" site is an interactive one for teachers and students. It provides instruction in writing lab reports and offers examples of such reports: http://labwrite.ncsu.edu

FIGURE 4.9
Laboratory Report Rubric

CRITERIA	EXEMPLARY	ADEQUATE	RESUBMIT
PROBLEM STATEMENT	**5 points** Problem is stated as a question. Problem is appropriate to assigned task.	**3 points** Problem is clearly identified but may not be in the form of a question, and problem is appropriate to assigned task.	**1 point** Inappropriate problem or problem that is insufficiently identified/explained.
RELEVANT SCIENTIFIC CONCEPTS	**7 points** Student explains all scientific concepts (correctly using scientific terms/vocabulary) relevant to the problem, explaining why these are relevant. Student provides references for the background information from text or reference materials.	**5 points** Student explains several scientific concepts (correctly using scientific terms/vocabulary) relevant to the problem, but may fail to include others. Student explains why the concepts presented are relevant, and student provides references from text or reference materials.	**2 points** Scientific concepts are missing *OR* irrelevant to problem *OR* student fails to correctly use scientific terms/vocabulary *OR* student fails to provide references from text or other reference materials.

CRITERIA	EXEMPLARY	ADEQUATE	RESUBMIT
HYPOTHESIS	**5 points** Hypothesis/prediction is based on proper use and interpretation of background information and is not just a guess. Hypothesis is clearly stated and aligns with the experiment (the hypothesis can be supported or refuted by the data collected).	**3 points** Hypothesis/prediction is based on background information, but some essential information may be missing or unclear. Hypothesis is aligned to experiment.	**1 point** Hypothesis/prediction is not supported by background information *OR* is not aligned to the experiment.
MATERIALS/ EQUIPMENT NEEDED	**5 points** All materials and equipment needed are listed. Illustrations/diagrams are provided to enhance clarity. Such graphics include labels giving names of relevant equipment/materials.	**3 points** All materials and equipment needed are listed.	**1 point** List of materials/ equipment is incomplete *OR* missing.
EXPERIMENTAL PROCEDURES	**7 points** Steps in procedure are numbered and sequenced logically. Sufficient explicit details are included to allow the reader to duplicate the experiment.	**5 points** Steps in procedure are sequenced logically. Sufficient explicit details are included to allow the reader to duplicate the experiment.	**2 points** Steps are not sequenced logically *OR* insufficient details are present to allow the reader to duplicate the experiment.
SAFETY PRECAUTIONS	**5 points** All possible relevant safety issues are listed, and student describes how to avoid/minimize such issues. There is little or no risk of safety or discomfort to the student.	**3 points** The most important and relevant safety issues are listed, and student describes how to avoid/minimize these issues. However, student fails to mention some minor safety issues that are always relevant (as checking glassware for cracks). Some further explanations could minimize discomfort to the student, but safety risk is minimal.	**1 point** Important safety issues are not listed *OR* student fails to describe how listed issues can be avoided/minimized. Safety risk is not appropriate.
RECORDED DATA	**10 points** Student identifies controls, dependent, and independent variables correctly. Data is displayed graphically in an easy to understand form. If measurements are taken, correct units are included. All significant data (data needed to solve the problem) is recorded.	**7 points** Student identifies controls, dependent, and independent variables correctly. If measurements are taken, correct units are included. All significant data is recorded.	**4 points** Student fails to correctly identify controls or variables *OR* no units are included *OR* not all significant data is recorded.

FIGURE 4.9 (continued)

CRITERIA	EXEMPLARY	ADEQUATE	RESUBMIT
CALCULATIONS	**7 points** Formulas are provided and used correctly (e.g., Density = Mass divided by Volume). All measurements contain proper units. All calculations are performed correctly, yielding accurate answers.	**5 points** Appropriate formulas are used, but may not be written out. All measurements contain proper units. Most calculations are performed correctly, yielding accurate answers (inaccuracy must not occur in greater than 2% of calculations).	**2 points** Formulas are inappropriate or not in use *OR* measurements lack units *OR* greater than 2% error on calculations.
CONCLUSION (WORDS)	**15 points** In the conclusion, student states if hypothesis is accepted or rejected. Any conclusion is based upon the data collected and the rationale for the conclusion is completely explained by referring to this data. The student's rationale leads correctly to the acceptance or rejection of the hypothesis.	**10 points** In the conclusion, student states if hypothesis is accepted or rejected. The conclusion, while supported by the data, may be incompletely explained.	**5 points** Conclusion does not refer to the hypothesis *OR* conclusion is not supported by the data.
CONCLUSION (GRAPHIC)	**7 points** Graph form chosen is appropriate for the data. All data points are accurately plotted. Graph contains an overall title and axes are labeled. Intervals are appropriate and consistent and measurements are given. Graph is easy to read and understand.	**5 points** Graph form chosen is appropriate for the data. All data points are accurately plotted. Graph contains an overall title and axes are labeled. Intervals are consistent and measurements given. Some use of inappropriate intervals may reduce clarity, making graph hard to read.	**2 points** Graph is missing *OR* form is inappropriate *OR* required titles/labels are missing *OR* intervals are not consistent *OR* measurements are not given.
REFLECTION	**5 points** Student uses the rubric to rate his/her performance and highlights areas of weakness. Student suggests appropriate strategies for addressing these weaknesses.	**3 points** Student uses the rubric to rate his/her performance and highlights areas of weakness.	**1 point** Student fails to use rubric to rate his/her performance *OR* uses rubric, but fails to highlight areas of weakness.

Score Conversion

60–75 → A   46–59 → B   34–45 → C   23–33 → D   Below 22 → Resubmit

CHAPTER FIVE

Analyzing Assessment Data and Making Instructional Modifications

Analyze Assessment Data

Modify Instruction

So far in this publication, we have explored two quadrants of the Classroom Assessment Cycle. These are: 1) Clarify learning targets and 2) Gather evidence in a variety of ways. These two quadrants compose the half of the Classroom Assessment Cycle that is most commonly implemented in classrooms. The remaining two quadrants of the Cycle are less frequently seen in classrooms. It seems that teachers sometimes stop moving through the Cycle after they collect assessment data. They fail to follow through and analyze the data in order to make inferences or draw conclusions about student learning. Let's examine transcripts of two interviews with science teachers to see which of the two teachers is completing the Cycle. Both teachers teach third grade.

Interview One: Interviewer (I) and Mrs. Esposito (E)

I: I see in your unit plan that you've included a diagnostic assessment on Day One. Can you tell me about this assessment?

E: Yes. The students will start a KWL chart (What I Know, What I Want to Know, What I Learned) before we begin our study of plants. They'll keep this in their notebooks and add to it every day.

I: What is your goal for this activity?

E: I want to activate the prior knowledge of the students and then encourage them to track their own learning.

I: How will you know if you've met these goals?

E: They'll turn in their completed KWL charts at the end of the unit.

I: How will your assessment data be used to make instructional modifications?

E: If some students turn in incomplete KWL charts at the end of this unit, I'll conference with these students.

Interview Two: Interviewer (I) and Mrs. Gregory (G)

I: Tell me about this seminar you're planning on scientific inquiry.

G: The students will do some text reading and define terms (e.g., problem, hypothesis, control) before the seminar. Then, I'll present a scientific problem to them about plants. For example, I might just refer to the plant we have in the classroom that looks like it's dying and ask, "What can we do?" This will begin the class discussion on the scientific inquiry.

I: What is your goal for this activity?

G: Actually, I have two. I want to build student knowledge of steps involved in a scientific investigation, and I want them to practice their seminar skills (listening, speaking, critical thinking).

I: How will you know if you've met these goals?

G: I'll take anecdotal notes during the seminar, using a matrix I've prepared. This will help me keep track of the seminar skills demonstrated by each student. To check their knowledge of scientific inquiry, I'll give the students another problem and ask them to design an experiment (working in small groups) that will help solve the problem.

I: How will your assessment data be used to make instructional modifications?

G: I'll review my notes about the seminar and conference with students to give them feedback about their performances. In the individual conferences, we'll brainstorm strategies for improving their performances. Students will present their ideas for the new experiment to the class and will receive feedback from me and from the other student groups. Based on the presentations, I will note areas of difficulty and plan either some learning center activities or some re-teaching strategies to use with small groups in order to address these.

Both teachers are collecting assessment data. It is clear that Mrs. Gregory is utilizing the data to impact her instruction in a more timely manner than that employed by Mrs. Esposito. For example, if Mrs. Esposito only takes up her diagnostic KWL chart at the end of the unit, she is missing the opportunity to discover what her students knew about plants before beginning the unit. She is also failing to identify possible misconceptions the students hold. Without this knowledge, she cannot plan instruction that will effectively eradicate these misconceptions. Therefore, her "diagnostic" assessment is not actually being used to diagnose student learning. In fact, she is using this KWL chart more as a formative assessment. (She seems to be using the KWL more as a self-assessment tool for students; one that helps students understand what they've learned and what they still need to learn.) Looking at the KWL at the end of the unit means that Mrs. Esposito is not using assessment data to help her adjust instruction in *this* unit. The data she collects may help her structure future units, but the data arrives too late to help students learn the material in the current unit. So, no immediate instructional modifications are being made by Mrs. Esposito.

Conversely, Mrs. Gregory is using her assessment data to continuously modify her instruction. She is collecting anecdotal data on seminar performances, reviewing these, and then helping students strategize ways to improve. She is also using the assessment of the student experiment presentations to differentiate instruction for those who are struggling. She plans to note areas of difficulty and then address these in learning centers or by re-teaching small groups of students. In this manner, Mrs. Gregory's assessment data has immediate impact upon instruction. Students receive timely feedback that will help them improve performance in the current unit.

From these two scenarios, we see that it is not enough to collect the assessment data; the teacher must also:

- Review the assessment data immediately after collection.
- Interpret the data, formulate inferences and/or conclusions.
- Provide feedback (verbal, written comments, rubric scores, etc.) to students relative to their performances.
- Use the findings of the analysis to immediately adjust instruction.

In order to proceed effectively through these steps, we need to examine what is meant by analysis and look at ways we can modify instruction to enhance individual student performance.

Analysis

The graphic in **FIGURE 5.1** provides a perspective on the analysis of data. The organizing question that the teacher must ask before beginning this analysis is, "What did the student learn and how well did he/she learn it?" We have seen that it is important to ask this question frequently; perhaps even on a daily basis. In this manner, assessment data can have an immediate effect on instruction in that the teacher can intervene quickly to address individual student learning difficulties. Data analysis can provide teachers with insights about student learning in the short term. However, continuous data analysis can also be useful in the long term. As we examine **FIGURE 5.1**, let's work through an example of the long-term use of data analysis.

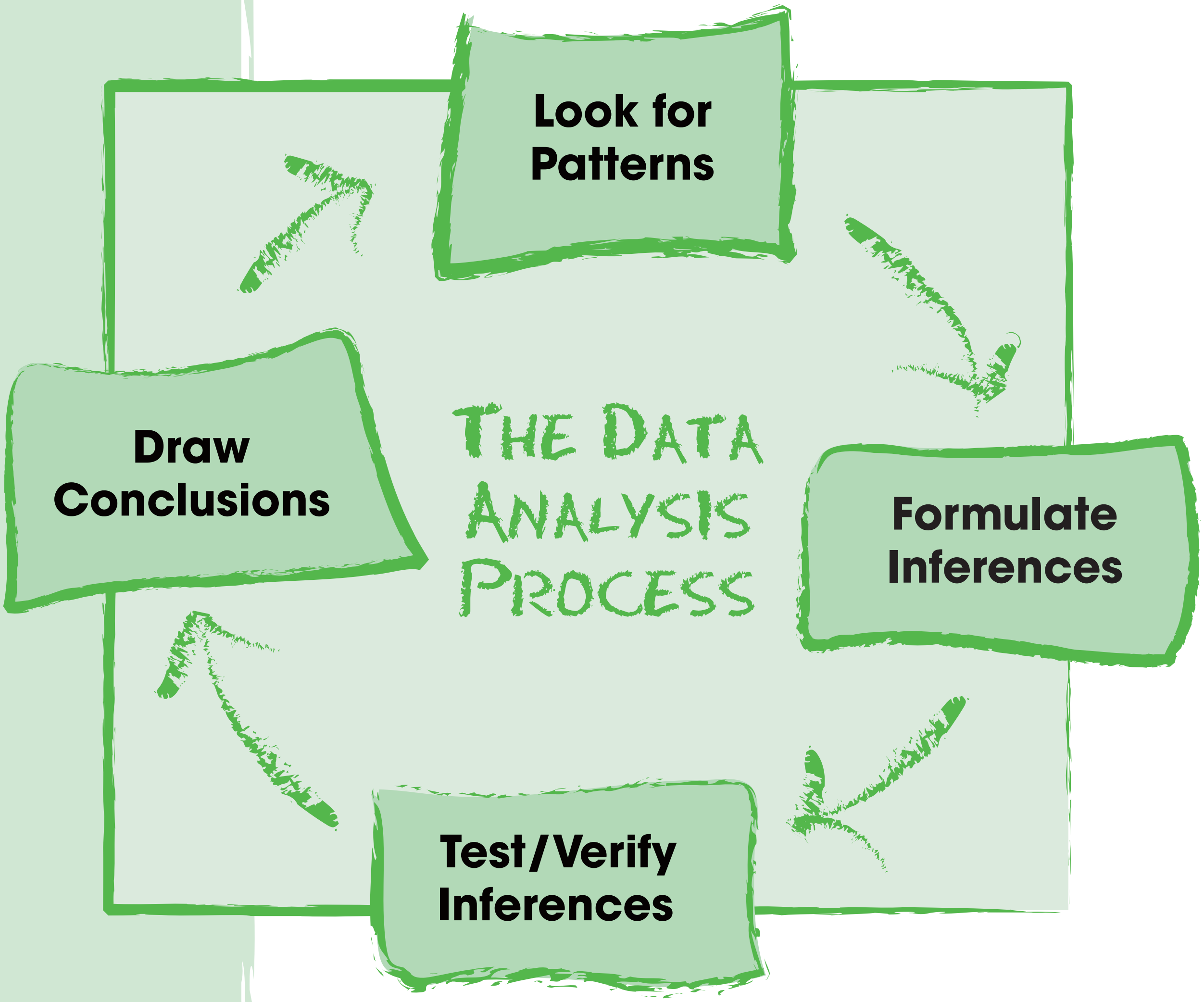


FIGURE 5.1
Data Analysis Process

At the end of the grading period, Mr. Gilley, a sixth-grade teacher, sat down to examine the assessment data he had collected over the quarter. He was particularly interested in reviewing assessments from those students who appeared to be struggling—those students who most often needed individual attention from him in order to be successful. He also looked at students who had low average grades. Surprisingly, one such student was Luci.

Before reviewing the assessments, Mr. Gilley had thought of Luci as one of his best students. She always seemed attentive and appeared to understand all the concepts presented. Looking at Luci's assessments in the Adaptations of Animals unit revealed the following:

1) Diagnostic paragraph that asked students to define "adaptation," give one example of an animal adaptation, and explain how this adaptation aided the animal to survive: Luci did not attempt to define adaptation, but wrote, "Polar bears—fur."

2) Reading comprehension sheet that included fill in the blank and short-answer question. This worksheet followed an assigned text reading on animal adaptations: Luci successfully completed the fill-in-the-blank questions. On the short-answer questions, Luci only attempted two of the four questions. She wrote her answers as phrases, not as complete sentences. Most of her answers were factual in nature, even when the question asked for analysis. This caused her to receive a low assessment on this assignment.

3) Matching test on terms from the chapter: Luci made a 100.

4) Animal Adaptation Project in which students were assigned a particular climatic zone and then asked to create an imaginary creature who could successfully survive in this zone. The project required that students draw their creature and then write a supporting paragraph identifying the adaptations and explaining why these adaptations would make the creature successful: The rubric used for scoring this project is shown in Figure 5.2. The rubric shows that Luci provided a detailed drawing of her creature and labeled this illustration with arrows and words as "claws—catch food." Her animal was clearly well adapted to survive in the tundra zone she was assigned. Luci's paragraph was very sketchy. It contained a list of adaptations with few explanations. Therefore, she earned very few points within the "paragraph" portion of the rubric. This negatively impacted her overall performance on this project.

FIGURE 5.2
Animal Adaptation Project Rubric—Luci's Scores

CRITERIA	EXCELLENT	ADEQUATE	RESUBMIT
DRAWING SIZE	**2 pts.** Animal is sized to use most of the space on a standard piece of notebook paper	**1 pt.** Animal is sized to use at least ¾ of the space on a standard piece of notebook paper	**0 pts.** Animal takes up less than ¾ of the space on a standard piece of notebook paper
DRAWING DETAILS	**12 pts.** Animal is shown with more than 3 adaptations. Each adaptation is labeled with a name and a function. The function matches the adaptation (Ex: teeth are labeled for fighting or eating functions, not for adapting to weather).	**9 pts.** Animal is shown with at least 3 adaptations. Each adaptation is labeled with a name and a function. The function matches the adaptation.	**0–8 pts.** Animal is shown with fewer than 3 adaptations *OR* adaptations are not labeled with names and function.
PARAGRAPH	**8 pts.** Paragraph accurately explains why animal will need the adaptations for the assigned environment. At least 4 adaptations are explained.	**6 pts.** Paragraph accurately explains why animal will need at least 3 adaptations for the assigned environment.	**0–5 pts.** 2 pts. Paragraph fails to accurately explain the animal's need for at least 3 adaptations. Only 2 adaptations listed, not completely explained
GRAMMAR	**5 pts.** Paragraph contains no spelling, capitalization, or other grammatical errors. Student writes in complete sentences.	**3 pts.** Paragraph contains 2–3 grammatical errors. Student writes in complete sentences.	**0–2 pts.** 1 pts. Paragraph contains more ~~than 3 grammatical errors~~ *OR* student fails to write in complete sentences.
ADAPTED TO ENVIRONMENT	**4 pts.** More than 3 adaptations match animal needs in the assigned environment.	**3 pts.** At least 3 adaptations match animal needs in the assigned environment.	**0–2 pts.** Less than 3 adaptations match animal needs in the assigned environment.

Score Conversion
28–31 → A 23–27 → B 18–22 → C
Below 18 → Student must resubmit assignment

Luci's Score: 18

APPLICATION

Before reading on, answer the following questions. Compare your answers to the ideas presented in the following section.

1. What patterns of student behavior did you discern?
2. What inferences did you draw about Luci's learning and/or performance?
3. Do you need more data? If so, what types of data would you like to collect?
4. What would you suggest as a "next step" in order to improve Luci's learning and/or performance?

After reviewing this assessment data, Mr. Gilley began to ***look for patterns****. One pattern that seemed particularly strong was Luci's performance on written assignments. It appeared that Luci rarely wrote in paragraphs. She appeared to use labels or phrases in lieu of complete sentences or organized, cohesive paragraphs. Another pattern that emerged was Luci's consistent high performance on selected response items. This behavior seemed more in line with Mr. Gilley's own perceptions of Luci's performance.*

Once he had noted patterns in Luci's performances, Mr. Gilley reflected on underlying causes for these patterns. He began to ***formulate inferences*** *about her academic behaviors. It was obvious from the examination of her selected response work that Luci was learning science concepts. It was also obvious that writing was a problem for Luci. Mr. Gilley tried to list possibilities for this problem:*

a) Luci has not had sufficient instruction in writing skills.

b) Luci has had little experience in writing.

c) Luci has a specific learning disability related to writing.

In order to narrow these possibilities, Mr. Gilley decided he needed more information. He would not be able to ***test or verify his inferences*** *without collecting additional data. First, he consulted Luci's cumulative folder to see if she had ever been tested for specific learning disabilities. He found that she had been tested, but that no specific learning disabilities were found. He then talked with Luci's current language arts teacher, Mrs. Riley, about her performance in that class. Mrs. Riley reported that she, too, had noticed Luci's reluctance to attempt, and poor performance on, written assignments. She told Mr. Gilley that she had begun to work closely with Luci, giving her individual attention and extra help in order to develop her writing skills. Finally, Mr. Gilley also conferenced with Luci and her parents and shared Luci's assessment results with them. This conference revealed that Luci felt unprepared to write and that she was reluctant to "show her ignorance" as she put it, by writing. She had had some bad experiences in the past, feeling ridiculed and belittled by some of the feedback she had received on her writing. Because she found writing to be risky to her self-esteem, Luci avoided it.*

Through this data collection process, Mr. Gilley obtained sufficient information to ***draw conclusions*** *about Luci's writing performance. He concluded that Luci's poor writing skills were the result of insufficient instruction and insufficient practice, rather than attributing this poor performance to a specific learning disability. Once he had identified the problem through the* ***data analysis process*** *described here, Mr. Gilley was ready to plan instructional modifications to enhance Luci's academic performance in his science class.*

Instructional Modifications

In the above data analysis example (Mr. Gilley and Luci), we saw that after the analysis process, Mr. Gilley's next step was to plan needed instructional modifications. By analyzing his students' (Luci's and all the other students') work and then designing instructional strategies to meet the needs of his students, Mr. Gilley demonstrated an example of *differentiation* in the classroom.

Differentiation

When teachers modify instructional plans to meet the needs of specific students, they differentiate instruction. Differentiated instruction simply means "teachers reacting responsively to a learner's needs" (Tomlinson & Allan, 2000, p. 5). Teachers who differentiate instruction know that one size does ***not*** fit all. They recognize that they teach a group of diverse learners with different learning styles, interests, and prior experiences.

Almost all teachers differentiate instruction to some extent. For example, when a lecturing teacher notices a puzzled expression on a student's face, she may pause, and then present the information again, but use a different manner of explanation this time. In this process, she is reacting to an individual student's needs and adjusting her instruction to fit that student's needs. In this chapter, we emphasize that teachers need to go beyond such stimulus-response reactions and think proactively about differentiation in order to make differentiation a natural part of the classroom environment.

In the past when we have urged teachers to differentiate in their classroom, we have heard a chorus of groans and muttered comments such as "Impossible!" and "Not enough time!" This led us to facilitate a discussion of differentiation to see if our definition matched that of the teachers. We found that it did not. Before we go further with a discussion of differentiation in this chapter, we need to clarify the term. We have already stated what differentiation IS (teachers responding to individual student needs). However, we will now explain what differentiation is ***not***.

Differentiation myths:

1) Differentiation is equivalent to individualized instruction.

"Individualized instruction" evokes images of students sitting with self-paced, self-checking booklets. In such classrooms, one envisions no peer interactions and few teacher/student interactions. Today, many home school students experience a new brand of individualized instruction—a one-on-one experience between the teacher and the student. We don't want to re-institute the first description of individualized instruction, and we aren't able to do the second. In differentiated instruction, we are ***not*** individualizing or tailoring instruction for each and every student.

2) *Differentiation is only important for low-ability students.*

High-ability students need differentiation as much as low-ability students or they become bored, restless, and disruptive.

3) *Differentiation means that we can never have whole-group activities.*

Everything in a differentiated classroom does not take place in small groups. Whole-group activities do occur. The teacher may demonstrate or model a skill or process—tell a story, share information, etc. When we hold students accountable for the learning acquired through whole-group activities, we may do this in a variety of ways.

4) *Differentiation requires ability grouping.*

Were you ever ability grouped? Didn't you figure it out, even though your group had a non-judgmental name (e.g. the Redbirds)? Students will figure it out, too. Differentiated instruction doesn't mean setting up the classroom to keep the low kids low and the high kids high—we don't want any self-fulfilling prophecies. Rather, a differentiated classroom takes the "Army" approach: Be all that you can be.

5) *Differentiation means the "bright kids" have to do more.*

No, we don't expect the kids who catch on the fastest to do ***more***—they may just do different things. They may be excused from some practice activities, as they've already "gotten" it—they don't need the practice.

6) *Differentiation means I have to incorporate choices for students in every assignment, activity, or assessment.*

Teachers should only differentiate instruction when there is a perceived need to do so. It is important to remember that teachers' classrooms aren't judged to be "differentiated" or "not differentiated" based on a single day or a single assignment. Differentiating means using a variety of instructional strategies and assessment methods as your students move through the curriculum.

Returning to the example of Mr. Gilley's classroom, Mr. Gilley used the data analysis process to determine the needs of his students. We know he found that Luci needed extra practice and extra support in writing skills. However, we don't have the results of the other analyses he performed. Perhaps he found other students in the class who shared Luci's need for instructional modifications in writing. He certainly found students with other learning needs. By pinpointing the learning strengths and weaknesses of the students in his class, Mr. Gilley was able to plan instruction to meet the needs of all his students. How might this be accomplished? One expert in differentiation, Carol Ann Tomlinson (1999), recommends planning for differentiation within three different dimensions of the classroom: Product, Process, and Content. The chart in **FIGURE 5.3** provides definitions of these dimensions and gives examples of how teachers may differentiate in these areas.

FIGURE 5.3
Dimensions of Differentation

DIMENSION		DEFINITION	IDEAS FROM THE EXPERTS
DIMENSION	CONTENT	Facts, concepts, generalization, principles, attitudes, and skills related to the subject. Includes both what the teacher plans for students to learn and how the student gains access to the desired knowledge, understanding, and skills.	Choosing text/novels/reading assignments to address different reading levels. Present info one time as "break it down from the whole"; next time build it up from its parts. Use readings, videos, computer demos, audio recordings to present material. Use pre-assessments as journal entries, class discussions, quizzes to assess student readiness to learn. Provide a note-taking matrix to some students.
DIMENSION	PROCESS	How the learner comes to make sense of, understand, and "own" key facts, concepts, generalizations and skills. (Synonym for Activity)	Encourage student conversations—have them explain concepts in "kid language." Say, "This reminds me of…" to help students make connections. Ask students to find patterns, categories, similarities/differences in data. Ask students to make and justify judgments. Work with students to develop rubrics. Allow for varied working arrangements (work alone or in groups).
DIMENSION	PRODUCTS	Items the student can use to demonstrate what she has come to know, understand, and be able to do as the result of study (e.g., portfolio, test, project).	Encourage students to express what they have learned in various ways. Provide or encourage varied use of resources in preparing products. Provide product assignments at varying degrees of difficulty. Use a wide variety of assessment methods. Encourage culminating assessments such as "Senior Projects."

For example, Mr. Gilley knows that Luci and several other students in the class have poor writing skills. He could choose to differentiate in the "Process" dimension by providing "choice as to how students will express knowledge (write a paper, make a diagram, construct a model)." This would allow Luci and the other students to express their scientific knowledge in ways other than writing. However, if he wishes to enhance the writing skills of his students, Mr. Gilley may choose to differentiate in the "Products" dimension by allowing for varied working arrangements (work alone or in groups). Here, he could team Luci or other students struggling with creating written products with more proficient peers or he could create sub-groups of struggling students within the classroom and spend more instructional time with them.

APPLICATION

Analyze your assessment data from the last grading period. Once you have identified student learning needs, list at least three ways you can differentiate within each of the three dimensions (Products, Processes, and Content).

When teachers make instructional modifications by differentiating in one of these three areas (Content, Products, Process), they are responding to student needs. Before they can differentiate, they must first identify these needs through a data analysis process. In Chapters Three and Four, we explored different assessment methods teachers can use and different venues for providing feedback to students. In this chapter, we have emphasized reviewing the assessment data and analyzing it to determine students' learning needs so that instructional modifications to meet those needs can be designed. With this chapter, we complete the Classroom Assessment Cycle. Our final chapter provides further examples, showing the complete implementation of this cycle in science classes. At the end of this manual, you will also find an Appendix that discusses ways of communicating assessment results to parents and students.

CHAPTER SIX

Implementing the Cycle

This manual has introduced the Classroom Assessment Cycle as a means of enhancing student achievement. The Cycle prescribes a process that teachers can use to align their classroom assessments with research-based best practices. By implementing all four quadrants of the Classroom Assessment Cycle, teachers can:

1) Clarify science learning targets for their students.
2) Gather assessment evidence in a variety of ways and provide meaningful feedback to students on their academic progress in science.
3) Analyze assessment data to determine the learning needs of the students in science classes.
4) Modify science instruction to meet those learning needs. (See **FIGURE 6.1**.)

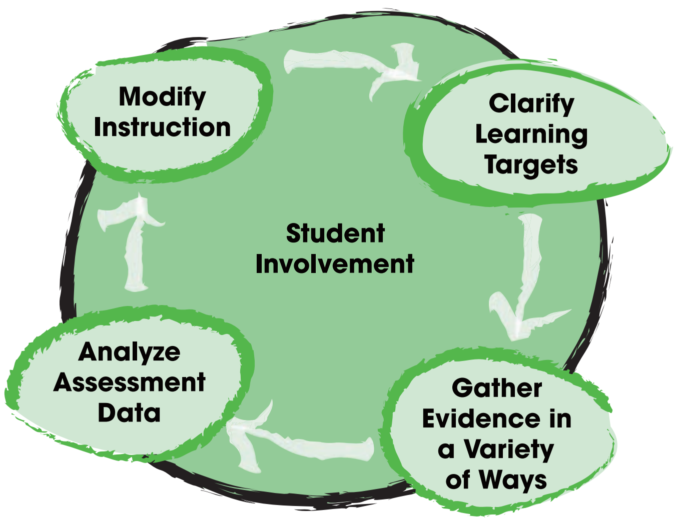


FIGURE 6.1 The Classroom Assessment Cycle

In this last chapter of the manual, we provide examples of science lessons that demonstrate how teachers at different levels may implement the Classroom Assessment Cycle. We think of these examples as snapshots from the classroom. They focus on only one day's lesson, rather than on a complete unit. The examples provide descriptions of previous activities, detailed plans for today's lesson, and a foreshadowing of tomorrow's science class. By studying these examples, teachers may obtain a clearer idea of how the Classroom Assessment Cycle works on a day-to-day basis in the science classroom. Teachers may then feel prepared to design lessons (and eventually whole units) using this assessment approach.

The first classroom snapshot is from Mr. Tighe's third-grade class. The chart in **FIGURE 6.2** categorizes the classroom activities into particular quadrants of the Classroom Assessment Cycle. This chart is followed by two **FIGURES** (**6.3** and **6.4**) that provide more information about the assessment and instruction in this classroom. Finally, **FIGURE 6.5** provides a summary of the learning targets, assessments, feedback, data analysis, and instructional modifications found in this third-grade lesson.

FIGURE 6.2

Elementary Level Implementation of the Classroom Assessment Cycle: Third Grade, Mr. Tighe's Class

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
UNIT GOALS: "As a result of activities in grades K–4, all students should develop understandings of ❑ The characteristics of organisms ❑ Life cycles of organisms ❑ Organisms and environment" (National Research Council, 1996, p. 127) In addition, under the guide to the "abilities necessary to do scientific inquiry" (Content Standard A), students will ❑ "Ask a question about objects, organisms, and events in the environment ❑ Use data to construct a reasonable explanation ❑ Communicate investigations and explanations" (National Research Council, 1996, p. 122)	**DIAGNOSTIC ASSESSMENT:** Yesterday, Mr. Tighe asked students to list examples of living and nonliving things found in the classroom. In a class discussion, the following list resulted: Living ❑ Mr. Tighe ❑ Me (student) ❑ Bug on windowsill ❑ Potted plant Nonliving ❑ Pencil ❑ Desk ❑ Blackboard ❑ Flag	**DIAGNOSTIC ASSESSMENT ANALYSIS:** Mr. Tighe concluded that students were able to differentiate living from nonliving things. No misconceptions were noted.	Mr. Tighe plans to build on student's knowledge of living and nonliving things. He will challenge the students to not only classify items as living or nonliving, but ask them to synthesize a list of characteristics of living things. He will prompt this discussion by bringing in some examples to stimulate student thinking.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
TODAY'S LESSON: **SPECIFIC OBJECTIVES** (kid-language): Students will ❑ Use our eyes, ears, noses, and touch to learn about living things. We will NOT use our mouths to taste anything today. ❑ Make a list of ways "living" things are different from "nonliving" things. ❑ Share our lists and our reasoning with the whole class. These specific objectives are aligned with the "characteristics of organisms" National Science Education Standard and with the "abilities necessary to do scientific inquiry" one.			
LEARNING TARGETS: ***Reasoning*** (Classifying objects as living, non-living; Synthesizing characteristics of living) ***Product*** (List of characteristics) ***Knowledge*** (Characteristics of living, nonliving things) ***Process Skills*** (Observing)			

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
Mr. Tighe orally shares today's specific objectives with students, using kid language, and writes the objectives on the blackboard.	**FORMATIVE ASSESSMENT:** Mr. Tighe writes yesterday's lists of living and nonliving things on the board and reads this list aloud to students. He asks students to think about this ***question***: How are living things different from nonliving things? He asks students to form their discussion groups (previously organized for other lessons) and discuss this ***question***. He reminds students to look for the star by their names (on the bulletin board, where members of the different discussion groups are listed) to find out who is today's reporter in the group. **FORMATIVE ASSESSMENT:** Mr. Tighe circulates and observes the group discussions. He collects information on a data collection sheet. (See **FIGURE 6.3**)		
	FORMATIVE ASSESSMENT: Student groups present reports (***constructed response—performance assessment***) of the discussion. As groups provide a characteristic of living things, Mr. Tighe writes each characteristic on the board.	Mr. Tighe sees that only six of the seven characteristics of life have been revealed in the group discussions. He knows that some groups thought of more characteristics than others.	Mr. Tighe wants to help all the students in the class be able to list and understand all seven characteristics of life. He will use concrete examples to help students make connections between the words (breathing, moving, etc.) and the concept of seven characteristics of life.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
	Mr. Tighe gives each discussion group a set of concrete examples (See **FIGURE 6.4**). These may be objects or pictures. He asks the students to decide which of the six characteristics on the board best matches the object/picture. He gives each student group a set of six cards labeled with the six identified characteristics. His directions to the students are to group the objects/pictures under the correct card. He also asks students to look for one missing characteristic and gives them a blank card for this characteristic. **FORMATIVE ASSESSMENT:** Again, Mr. Tighe circulates and ***observes*** groups. He asks ***questions*** to prompt student thinking as he circulates. He also responds to or initiates ***teacher/student dialogues*** for this same purpose. He particularly ***questions*** the groups about objects/pictures intended to be examples of the seventh characteristic, but classified by the students into other categories. Sample questions he asks: "How is this (object/picture) like the other ones in this category? How is it different? He also asks each group, "Did you find a new category?"	While circulating, Mr. Tighe sees that students are confused because they think some of the objects/pictures can fit into more than one category. The students are unable to progress until this issue is addressed. Mr. Tighe also notes that students are having difficulty coming up with the seventh characteristic. Many groups are putting the sunflowers under "movement." One group has the picture of children and horses under "eating/drinking" and another has this picture under "growth." Similarly, group three puts the drummer and girl picture under "growth" and group five puts this under "movement."	Mr. Tighe alerts the entire class to let them know that some objects/pictures fit under more than one category. He asks the class how this should be handled. One student suggests that they could create a drawing of any objects/pictures that might be used as examples of more than one characteristic (effectively making a duplicate). Mr. Tighe accepts this idea, stating that he thinks having a duplicate is a good idea; he volunteers to take a Polaroid picture of the object/picture for groups who would prefer this method over drawing. Mr. Tighe probes student thinking through his questioning activities. For those groups having the most difficulty, he uses a ***series of focusing questions*** to help them overcome obstacles. He spends ***more time*** with some groups than with others.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
	FORMATIVE ASSESSMENT: Mr. Tighe asks students to orally present (***constructed response—performance assessment***) their classifications to the class and explain why they classified particular objects under the headings. He facilitates a group discussion, leading to consensus on classification. This is student-driven and concludes without the seventh characteristic being discovered, as students simply place all the objects/pictures under one of the existing headings. Science time ends at this point.	Mr. Tighe knows that students still have not uncovered the seventh characteristic of life. He wants them to know this last characteristic (response to change) before beginning the next science lesson. One of the main objectives in the next lesson is to have students realize that for things to be classified as "living," they must display ***all*** seven characteristics of life. He plans to bring in examples and ask students to tell him if these objects are living or nonliving (e.g., a wind-up toy). If students say this is non-living, Mr. Tighe will challenge them with a question: "But, this toy does move. Isn't movement a characteristic of life?" Mr. Tighe cannot begin this new lesson, however, until students are cognizant of the "response to change" characteristic.	Since the examples given for the seventh characteristic of life, "response to change," were not effective in soliciting this category from students, Mr. Tighe formulates a new plan. He searches the Internet and finds several sites showing video clips of animals reacting to stimuli (e.g., a frog catching a fly by shooting out his tongue; a flock of ducks flying up all at once after hearing a loud sound; a dog tracking a scent). He decides to begin this lesson with a short class discussion. He formulates the following questions to ask the students: 1) Think of the last time you were frightened. What frightened you? What did you do? 2) If you heard a very loud noise right now, show me what you would do. 3) Have you ever burned yourself on something hot? Act this out for me. In answering the questions, Mr. Tighe hopes that one student will introduce the concept of cause and effect or "one action leading to another action." Mr. Tighe will build on this and introduce "response" as a term. He will try to get students to understand that in every case, they are responding to a change in their environment. He will then show the video clips and ask students to identify both the change that occurred and the response. Finally, he will go back to the three problematic pictures (intended to show this in yesterday's lesson) and ask students to identify changes and responses in them. He will add "respond to a change" to the characteristics of life list before beginning the next science lesson.

FIGURE 6.3 Mr. Tighe's Data Collection Form

DISCUSSION GROUP	GROUP WAS OBSERVED DISCUSSING ONE OR MORE OF THE FOLLOWING CHARACTERISTICS OF LIVING THINGS						
	Breathing	Moving	Reproducing	Eating/Drinking	Growing	Excreting	Responding to Change
1	X		X	X			
2	X	X		X	X		
3	X			X			
4	X	X		X		X	

FIGURE 6.4 List of Concrete Objects Distributed by Mr. Tighe to Student Discussion Groups

TO ILLUSTRATE...						
Breathing	Moving	Reproducing	Eating/Drinking	Growing	Excreting	Responding to Change
GI Joe Frogman doll with scuba gear	Live ladybug in a jar	Stuffed toy: mother elephant and baby elephant with linked trunks	Stuffed toy: mother cat with six attached, nursing baby cats	Picture of children being measured for height	Betsy Wetsy-type doll, with bottle	Picture of girl beside drummer; girl has hands over ears
Picture of child frosting window with his breath	Picture of a cheetah, running	Picture of mother holding infant	Picture of contestants at a pie-eating contest	Picture of boy whose clothes are too small (pants too short, shirt too tight)	Plastic dog "poop" from joke shop	Pictures of sunflowers pointing towards the sun (time lapse)
Picture of fish blowing bubbles underwater	Picture of children in a dance recital	Mother, father, boy, girl dolls	Betsy Wetsy-type doll with bottle	Pictures of quintuplets—at birth, at 1 year, at 7 years	Picture of dog urinating on man's leg	Picture of children in field with horse in summer beside picture of children in field with horse in winter (horse's coat is heavier, change in clothing for children)

FIGURE 6.5 Summary of Classroom Assessment Elements Found in Mr.Tighe's Third-Grade Science Class

LEARNING TARGETS	ASSESSMENTS	FEEDBACK TO STUDENTS	ASSESSMENT DATA ANALYZED	INSTRUCTIONAL MODIFICATIONS
REASONING, KNOWLEDGE, PRODUCT ❑ Classifying objects as living/nonliving	**DIAGNOSTIC ASSESSMENT** ❑ Produce list of living and nonliving things	Written and oral feedback in the form of the compiled list that is shared with students, both orally and on the board.	List of living and nonliving things—no misconceptions, students were able to differentiate.	No remediation necessary. Mr. Tighe builds on this foundational knowledge.
REASONING, KNOWLEDGE, PROCESS SKILLS ❑ Observing concrete examples of living things ❑ Synthesizing characteristics of living things	**FORMATIVE ASSESSMENTS** ❑ Questioning (How are living things different from nonliving things? How is this object/picture like the other ones in this category? How is it different? Did you find a new category?) ❑ Observations (Data collection form for group discussions; observations of student oral presentations) ❑ Teacher/Student Dialogues (Initiates conversations to prompt or focus student thinking)	Verbal Feedback from teacher Verbal Feedback from teacher and from peers Verbal feedback from teacher	Answers to questions given by students; Data collection form; Student oral presentation data	**CONTENT DIFFERENTIATION** ❑ Provides concrete examples to help students discover seventh characteristic of life and to aid students in learning/understanding the six characteristics that were uncovered in the original discussion. ❑ Breaks down content into small "chunks" by asking probing questions. ❑ Uses video clips to present information. **PROCESS DIFFERENTIATION** ❑ Asks students to find categories. ❑ Work in groups. ❑ Solves problem of "more than one category" with student input. ❑ Spends more time, gives more individual attention to some groups. **PRODUCT DIFFERENTIATION** ❑ Students express knowledge in a variety of ways (discussions, lists, oral presentations).

In Mr. Tighe's classroom, the primary formative feedback mechanism was verbal feedback. The following middle school example from Mrs. Nagura's seventh-grade science class (See **FIGURE 6.6**) also utilizes verbal feedback. However, in this example, written feedback in the form of quiz grades, comments on science journals, and rubrics is also present. Both examples incorporate diagnostic as well as formative assessments, use a variety of assessments and teaching methods, and emphasize the continuous analysis of assessment data to inform/modify instruction. The middle school example, like the elementary one, is followed by a summary chart of the learning targets, assessments, feedback, data analysis, and instructional modifications used (See **FIGURE 6.11**).

FIGURE 6.6

Middle School Level Implementation of the Classroom Assessment Cycle: Seventh Grade, Mrs. Nagura's Class.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
UNIT GOALS: "As a result of their activities in grades 5–8, all students should develop understanding of ❑ Structure and function in living systems ❑ Reproduction and heredity ❑ Regulation and behavior ❑ Populations and ecosystems ❑ Diversity and adaptations of organisms" (National Research Council, 1996, p. 155) In addition, under the guide to "abilities necessary to do scientific inquiry" (Content Standard A), students will: ❑ "Develop descriptions, explanations, predictions, and models using evidence ❑ Think critically and logically to make the relationships between evidence and explanations ❑ Communicate scientific procedures and explanations" (National Research Council, 1996, pp. 145–146)	**DIAGNOSTIC ASSESSMENT:** 1) ***Selected response assessment:*** Vocabulary quiz on terms from text reading, including: recessive, dominant, monohybrid cross, dihybrid cross, chromosome, gene. Mrs. Nagura marks the quizzes, indicating incorrect answers by circling them and then returns the papers to students.	**DIAGNOSTIC ASSESSMENT ANALYSIS:** The short-answer vocabulary quiz demonstrates that 22 of 26 students in the class can differentiate the associated terms (recessive from dominant, gene from chromosome, dihybrid from monohybrid crosses). The four students who had difficulty on the quiz tended to answer that chromosomes were on genes rather than conversely and could not identify mono- from dihybrid crosses.	Mrs. Nagura will begin class today by showing the mapped genes on a human chromosome from the Human Genome project (***teaching content in a new way***). She will stress that genes are small units of genetic information found on the chromosomes. She will then review mono- and dihybrid crosses by writing Punnet square examples (***modeling, demonstrating***) of each on the board (the class previously created Punnet squares for these two types of crosses in an earlier lesson). She will then give the students several gene pairs and ask them to create mono- and dihybrid crosses with these. Mrs. Nagura will circulate while students work on this, particularly ***visiting*** (and ***interacting*** with) those four students identified as having difficulty in this area.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
TODAY'S LESSON: SPECIFIC OBJECTIVES (kid-language): Students will: ❑ Use the genetic information given, create and compare the offspring of two parents ❑ Investigate these offspring for genetic diversity (Are they all alike? If not, how are they different? How many different offspring can you find?) ❑ Report your findings to the whole class ❑ Write an analysis of a cartoon in your science journal These specific objectives are aligned with the "reproduction and heredity" National Science Education Standard and with the "abilities necessary to do scientific inquiry" one. **LEARNING TARGETS:** ***Reasoning*** (develop explanations, think critically) ***Process skills*** (oral presentation) ***Knowledge*** (genetic information and its expression, the terms "genotype" and "phenotype") ***Product*** (accurate model of offspring, journal entry)	2) ***Constructed response—product assessment:*** Journal entry. Students reflect on the meaning of the cartoon. Explain this cartoon to someone who may not "get" it. (Cartoon shows a mom, dad, boy, and baby around the dining room table. The mom, dad, and baby all look alike, with sharp pointed chins, long, thin noses, wide lips, large ears, and round eyes. The boy does not have these characteristics. The caption of the cartoon shows the boy saying, "Mom, Dad, I failed genetics!"). Mrs. Nagura distributes the Genetics Cartoon Analysis Rubric (See **FIGURE 6.7**) for students to use. She scores the cartoon journal entry with this rubric so that students can identify strengths and weaknesses in their performances.	Mrs. Nagura read through the journal entries to see if students uncovered the deeper meaning of the cartoon. She wants students to observe that the boy does not appear to share the same physical characteristics seen in the parents and the baby. She hopes students will realize that this is a natural occurrence and that this is NOT an example of "failing" genetics. She also would like to see some discussion of the mechanisms that create genetic diversity. After reading the journals, Mrs. Nagura decides that most students didn't "get" the message. Most journal entries seemed to explain the cartoon by saying that the boy has failed his genetics class in school. Most students did not mention the different physical characteristics seen within the family. Only six students noticed this, and none of the six used recessive, dominant, linked, or unlinked genes (terms from the genetics chapter) in their explanations. They were more likely to simply state that "the boy doesn't look like his parents or the baby."	The analysis of the cartoon shows Mrs. Nagura that students do not yet have a deep understanding of how genetic diversity arises. She decides to use the cartoon again (as an ending reflection) after today's lesson to see if students' explanations about the cartoon change before planning any instructional modifications.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
Mrs. Nagura writes today's specific objectives on the board and reviews these with the students. She uses "kid language" for the written objectives and in her explanations of these objectives.	Mrs. Nagura asks students to sit with their lab partners to form groups of two. She passes out the direction sheet for the activity and reviews this with the students. (See "Creating a Baby Creature" in **FIGURE 6.8**). **FORMATIVE ASSESSMENT:** Mrs. Nagura checks comprehension of the instructions by asking (questioning) students to express these in their own words. She checks to see if students understand the code by writing several codes on the board and asking (questioning) students what the baby having these codes would look like. She shows students models of the two parent Creatures, so they can see what an assembled Creature looks like. She quizzes students (questioning) on the materials to use in creating particular physical characteristics.	Through ***questioning***, Mrs. Nagura finds that students are able to translate the genetic code into physical characteristics. They know what materials to use to create these characteristics.	No instructional modifications were needed for this portion of the lesson, based upon the analysis of questioning data. However, Mrs. Nagura circulates and observes (***one-on-one interactions***) the student groups while they make their Creatures. She checks with several groups to see if they are translating genotypes into phenotypes correctly.
	Once all the Creatures have been placed in the playpen, Mrs. Nagura passes out the Discussion Sheet (See **FIGURE 6.9**). Student groups will discuss the questions on this sheet. One student will record the answers and the other student will report the answers to the class. The class may ask questions or make comments about the reports. Mrs. Nagura writes anecdotal notes while students are reporting. **FORMATIVE ASSESSMENT:** Students record answers on the discussion sheet (***constructed response-product assessment***) and present their findings to the class (***constructed response-performance assessment***).	Mrs. Nagura listens to the group reports and to the student comments. From this information, she concludes: 1) Students found no two baby Creatures that were exactly alike. Students were impressed with the level of diversity found. 2) Most student groups reported that the number of phenotypes would equal the number of genotypes, which is inaccurate. 3) Students found that the chromosomes they discarded would have made a baby that looked different from the one they actually constructed.	Mrs. Nagura brings the class into a whole-group session. She displays an overhead transparency of the decoding chart found in Figure 6.8 and draws students' attention to those genotypes that code for the same phenotype (DD and Dd; EE and Ee; FF and Ff; GG and Gg). She asks students to count the number of different genotypes listed on this chart (there are 21). She then asks them to count the number of different phenotypes (there are 17) on this chart (***teaching content in a new way; asking students to find patterns***). Mrs. Nagura then refers back to question 4 on the Discussion sheet. "In this case, are there more genotypes possible, more phenotypes, or the same?" she asks. Since students

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
	Students receive verbal feedback from the teacher and from peers on their answers. This feedback is also self-assessment, as students can check their own answers against those from other groups. Hearing others' ideas can stimulate deeper thinking.	4) Students explained that Creatures only have 14 chromosomes and that they get one from each parent. So, the reduction is needed to be sure that baby Creatures don't get too many chromosomes. 5) There appears to be controversy on the linked chromosomes question—student answers varied greatly and students had a tough time justifying their answers to others. These observations tell Mrs. Nagura that she must find ways to help students understand 1) that if two different gene combinations result in the same physical trait (as EE and Ee both result in two eyes), then the number of phenotypes will be less than the number of genotypes 2) how linked genes decrease the potential for genetic diversity	just counted them, they are able to answer that there are more genotypes than phenotypes. Mrs. Nagura asks a student to put what they just observed in his own words (***using "kid language"*** and ***actively involving students in the learning process***). He answers, "Because two different genotypes make the Creature look the same way, there are less phenotypes than genotypes." To help students understand the linked genes, Mrs. Nagura ***demonstrates*** two sets of genes—linked and unlinked. See **FIGURE 6.10** to see the figures she presents. She shows the linked genes physically joined to one another (***using a concrete model***), so that A can only be with B and "a" can only be with "b" (can't have aB or Ab). She then works through the Punnet Squares of the two sample (linked and unlinked) crosses (***modeling***). She lists the different phenotypes that would arise in each type and cross and shows (***use of concrete example, demonstration***) students how linkages between genes cuts down on the number of phenotypes possible.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
	Mrs. Nagura displays the mother and father Creatures to the class. She asks the class to compare the children to these parents. **FORMATIVE ASSESSMENT:** Mrs. Nagura asks (***questioning***) students, "Is it possible for two parents to have a child that does not look exactly like either parent? Is it possible for the same two parents to have another child, again that does not look exactly like either parent and also doesn't look exactly like the first child?" She then asks, "Based upon your observations in this activity, would you say having two children who both look different from their parents and different from each other is OK or did something go wrong? Is it OK for one child to look similar (but not exactly like) one or both parents and the other to look nothing like either parent? Would this be a natural occurrence, or would something have to go wrong for this to occur?" Once students discuss these questions ***(verbal feedback from peers and teacher***) and arrive at a consensus (both situations are "OK" and could naturally occur, just as they did with the Creatures), Mrs. Nagura gives the final assignment of the day. She puts up the same cartoon from yesterday and asks students to observe the family closely and then to write (constructed response-product assessment) an analysis of the cartoon in their journals. Again, students receive feedback via the Genetics Cartoon Analysis Rubric. Mrs. Nagura reminds students to examine the assessment of their first analysis before beginning to write this one. She encourages them to use the rubric comments to improve their performance.	Mrs. Nagura reads the new (second) journal entries about the cartoon. She finds that all but 3 students now perceive that the boy does not look like his parents or the baby. All students except these three are able to articulate that this is a natural occurrence. 75% of the students use the term "unlinked genes" to help explain the diversity that might arise. Mrs. Nagura feels that the majority of the class now has a better understanding of how linkages between genes reduce the genetic diversity.	Tomorrow, Mrs. Nagura plans to introduce a sex-linked trait, colorblindness. She will have the students list all the possible genotypes and phenotypes for offspring, given the genotypes of the parents. She will circulate while students work on this problem, and she will provide ***individual assistance*** to those students who either did not "get" yesterday's cartoon or did not use "unlinked gene" as part of their explanations of the cartoon. By using probing ***questions***, she hopes to lead students to an understanding that linkages reduce diversity of phenotypes. *This lesson was adapted from ideas found in: Soderberg, P. (November, 1992). Marshmallow meiosis.* The Science Teacher, *pp. 28–31.*

FIGURE 6.7

Genetics Cartoon Analysis Rubric

CRITERIA	EXCELLENT	ADEQUATE	POOR
SUMMARY	Journal entry accurately summarizes the event taking place in the cartoon, including all key elements so that a reader could create a "mind picture" of the cartoon from the summary.	Entry summarizes the event in the cartoon, but some details are missing/inaccurate. Reader could get a partial "mind picture" of the cartoon.	Summary is missing or so incomplete that a reader would have trouble creating a "mind picture" of the cartoon.
INTERPRETATION	Journal entry identifies symbols, metaphors, and/or representations used in the cartoon and explains what these objects stand for. Entry explains why the cartoon's creator chose these particular symbols, metaphors, or representation.	Journal entry identifies symbols, metaphors, and representations used by the creator, but fails to explain what these stand for. Entry attempts to explain why these objects were chosen, but explanation is occasionally inaccurate or incomplete.	Interpretation of symbols is missing or incomplete. Entry does not attempt to explain why the creator included particular objects in the cartoon.
ANALYSIS	The entry correctly identifies the underlying meaning of the cartoon and provides at least three supporting details from the cartoon to support this meaning. The explanation of the meaning accurately uses scientific terms from the chapter on genetics in the text.	The entry identifies a possible meaning, but provides only 1–2 supporting details.	Entry lists a possible meaning, but does not give any supporting details to support this meaning.

FIGURE 6.8
Creating a Baby Creature

Names of Group Members: ____________________

1) Each student group will receive a packet of chromosomes. The yellow chromosomes belong to Mom Creature and the black chromosomes belong to Dad Creature.

2) Notice that the chromosomes appear to be of different lengths and that each chromosome is labeled with a letter.

3) Lay all the chromosomes face down (so the letter is hidden) on the table. Sort them into same size groups (each same size group should contain four chromosomes (two yellow and two black).

4) Choose one yellow chromosome and one black chromosome from each same size group. Leave the chosen chromosomes on the table. Put the remainder of the chromosomes back in the packet and put these aside.

5) Chromosomal analysis on the Creatures has shown that each chromosome contains only one gene. This one gene codes for a particular trait. The attached chart shows you the genetic code (genotype) for each trait beside the physical description (phenotype) of this trait. You will use this code to create your Creature offspring:

GENOTYPE	PHENOTYPE	MATERIALS USED TO CREATE THIS TRAIT
AA	1 antenna	1 small nail
Aa	2 antenna	2 small nails
aa	No antenna	
BB	1 yellow hump on back	1 yellow mini-marshmallow
Bb	2 yellow humps on back	2 yellow mini-marshmallows
bb	3 yellow humps on back	3 yellow mini-marshmallows
CC	Green nose	1 green mini-marshmallow
Cc	Yellow nose	1 yellow mini-marshmallow
cc	Pink nose	1 pink mini-marshmallow
DD or Dd	Curly tail	1 pipe cleaner, curled
dd	Straight tail	1 pipe cleaner, not curled
EE or Ee	2 eyes	2 thumbtacks
ee	3 eyes	3 thumbtacks
FF or Ff	4 blue legs	4 blue push pins
ff	4 red legs	4 red push pins
GG or Gg	3 body segments	3 large marshmallows
gg	2 body segments	2 large marshmallows

6) Turn your mixed black and yellow chromosomes over and write down the codes you find here.

7) Use the chart above to convert the genotype (the two letters) into the phenotype (what the baby will look like). Gather the materials you need and assemble your baby Creature. You may use some toothpicks to hold your baby together. The toothpicks should be hidden, not visible when your baby Creature is finished. (Think of the toothpicks as bones.)

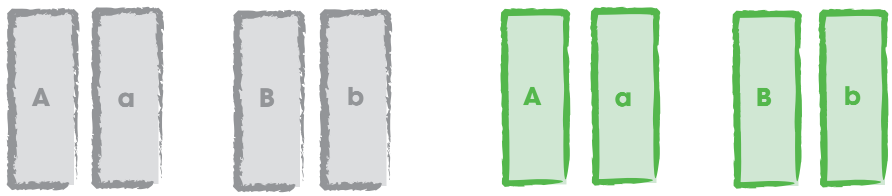
8) Write the names of the two members of your group on a piece of paper. Put this piece of paper in the Creature playpen (at the front of the room on the teacher's desk) and place your baby Creature on this paper. This piece of paper serves to identify your baby Creature.

FIGURE 6.9
Discussion Sheet for Creating a Baby Creature

1. How many baby Creatures were constructed in today's class?
2. How many were exactly alike (identical twins)?
3. How many were different (in at least one way) from all the others?
4. If there are 21 different genotypes possible for the baby Creature, are there the same number, more, or fewer phenotypes that are possible? Explain your answer.
5. You started this activity with 28 chromosomes (14 from Mom and 14 from Dad). You discarded 14 chromosomes (7 from Mom and 7 from Dad). Look at the chromosomes you discarded. If you had made your baby Creature using these chromosomes, would it have looked the same as the one you actually made? Explain your answer.
6. Why was it necessary to reduce the number of chromosomes (from the original 28 chromosomes to only 14) before creating your baby Creature? Explain. (In genetics, this "reducing" process is known as meiosis.)
7. Select the A (A or a) yellow chromosomes you used to create your baby Creature. Tape the yellow B (B or b) to the bottom of this chromosome. Do the same to your A (A or a) and B (B or b) black chromosomes. The A and G chromosomes are now linked together. How might this affect the genetic diversity? (Increase it? Decrease it? Why?)

FIGURE 6.10
Linked and Unlinked Genes

Mom's Genes on Chromosomes | **Dad's Genes on Chromosomes**



Mom's Possible Genotypes

AB
Ab
aB
ab

Dad's's Possible Genotypes

AB
Ab
aB
ab

PUNNET SQUARE OF DIHYBRID CROSS—UNLINKED GENES

		DAD'S POSSIBLE GENOTYPES			
		AB	Ab	aB	ab
MOM'S POSSIBLE GENOTYPES	AB	1) AABB	2) AABb	3) AaBB	4) AaBb
	Ab	2) AABb	5) AAbb	4) AaBb	6) Aabb
	aB	3) AaBB	4) AaBb	7) aaBB	8) aaBb
	ab	4) AaBb	6) Aabb	8) aaBb	9) aabb

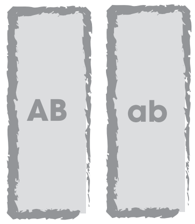
Offspring's genotypes are shown in shaded squares. Numbers indicate different phenotypes shown below. If same number appears beside two genotypes, they code for the same phenotype.

9 Possible Phenotypes for Baby Creature

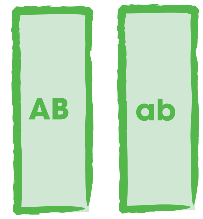
1) Baby's genotype is AABB, so baby has 1 antenna, 1 yellow hump
2) Baby's genotype is AABb, so baby has 1 antenna, 2 yellow humps
3) Baby's genotype is AaBB, so baby has 2 antenna, 1 yellow hump
4) Baby's genotype is AaBb, so baby has 2 antenna, 2 yellow humps
5) Baby's genotype is AAbb, so baby has 1 antenna, 3 yellow humps
6) Baby's genotype is Aabb, so baby has 2 antenna, 3 yellow humps
7) Baby's genotype is aaBB, so baby has no antenna, 1 yellow hump
8) Baby's genotype is aaBb, so baby has no antenna, 2 yellow humps
9) Baby's genotype is aabb, so baby has no antenna, 3 yellow humps

FIGURE 6.10
Linked and Unlinked Genes (continued)

Mom's Genes on Chromosomes



Dad's Genes on Chromosomes



Mom's Possible Genotypes

AB
ab

Dad's's Possible Genotypes

AB
ab

PUNNET SQUARE OF DIHYBRID CROSS—LINKED GENES

		DAD'S POSSIBLE GENOTYPES	
		AB	ab
MOM'S POSSIBLE GENOTYPES	AB	1) AABB	2) AaBb
	ab	2) AaBb	3) aabb

Offspring's genotypes are shown in shaded squares. Numbers indicate different phenotypes shown below. If same number appears beside two genotypes, they code for the same phenotype.

3 Possible Phenotypes for Baby Creature

1) Baby's genotype is AABB, so baby has 1 antenna, 1 yellow hump
2) Baby's genotype is AaBb, so baby has 2 antenna, 2 yellow humps
2) Baby's genotype is aabb, so baby has no antenna, 3 yellow humps

FIGURE 6.11

Summary of Classroom Assessment Elements Found in Mrs. Nagura's Seventh-Grade Science Class

LEARNING TARGETS	ASSESSMENTS	FEEDBACK TO STUDENTS	ASSESSMENT DATA ANALYZED	INSTRUCTIONAL MODIFICATIONS
REASONING, KNOWLEDGE, PRODUCT ❑ Demonstrate knowledge of scientific terms on vocabulary quiz ❑ Analyze genetics cartoon	**DIAGNOSTIC ASSESSMENT** ❑ Vocabulary quiz ❑ Cartoon analysis	Written feedback on quiz: wrong answers circled Written feedback on cartoon, using Genetics Cartoon Analysis Rubric	Answers to questions on vocabulary quiz; student analyses of cartoon	**CONTENT DIFFERENTIATION** ❑ Teach content in a new way, use pictorial content (visual from Human Genome Project) ❑ Model, demonstrate **PROCESS DIFFERENTIATION** ❑ Individualized attention (visiting and interacting with students while circulating)
PRODUCT ❑ Create a model of the baby Creature	**FORMATIVE ASSESSMENTS** ❑ Questioning (Asks comprehension questions about directions and materials to use and about translating the code into physical characteristics)	Verbal feedback from teacher	Answers to questions given by students	**CONTENT DIFFERENTIATION** ❑ Provides concrete examples to help see what a model looks like (shows Mom and Dad Creatures) ❑ Breaks down content into small "chunks" by asking comprehension questions **PROCESS DIFFERENTIATION** ❑ Asks students for find patterns (practice in translating the code) ❑ Asks students to express ideas in own words (kid language) ❑ Individualized attention (visiting and interacting with students while circulating)

LEARNING TARGETS	ASSESSMENTS	FEEDBACK TO STUDENTS	ASSESSMENT DATA ANALYZED	INSTRUCTIONAL MODIFICATIONS
REASONING, PROCESS SKILLS ❑ Develop explanations, think critically (Discussion Sheet) ❑ Make oral reports of answers to discussion questions	**FORMATIVE ASSESSMENTS** ❑ Constructed response—product assessment (written questions on Discussion Sheet) ❑ Constructed response—performance assessment (oral reporting) ❑ Teacher Observation (observes, listens to student performances during oral reporting)	Verbal feedback from teacher and peers	Answers to questions given by students; anecdotal notes taken by Mrs. Nagura	**CONTENT DIFFERENTIATION** ❑ Teaching content in a new way, asking students to find patterns (examining genotypes that code for same phenotype) ❑ Model, demonstrate (works through sample Punnet Squares) ❑ Provides concrete model (linked genes) **PROCESS DIFFERENTIATION** ❑ Using "kid language" and actively involving students (having students put ideas in their own words) ❑ Students express knowledge in a variety of ways (class discussions, written answers to Discussion Sheet questions, oral presentations)

LEARNING TARGETS	ASSESSMENTS	FEEDBACK TO STUDENTS	ASSESSMENT DATA ANALYZED	INSTRUCTIONAL MODIFICATIONS
REASONING ❑ Genetics Cartoon analysis	**FORMATIVE ASSESSMENTS** ❑ Questioning ("Is it possible for two parents to have a child that does not look exactly like either parent? Is it possible for the same two parents to have another child, again that does not look exactly like either parent and also doesn't look exactly like the first child?", etc. ❑ Constructed response—product assessment (analysis paragraph written in journal) ❑ Self-assessment (examine own past performance using rubric scores, plan how to improve scores in this assignment)	Verbal feedback from peers and teacher Genetics Cartoon Analysis Rubric	Second analyses of cartoon	**PROCESS DIFFERENTIATION** ❑ Ask students to find patterns, categories (list all possible genotypes and phenotypes) ❑ Individualized attention

The last exemplar in this manual provides a snapshot of a high school Biology class (**SEE FIGURE 6.12**). Mrs. Green is the teacher. Her students have been working on a problem-based unit concerning conjoined twins. In the example below, rubrics are the primary feedback mechanisms. As with the other samples, this high school snapshot ends with a summary of the learning targets, assessments, feedback, data analysis, and instructional modifications used (**SEE FIGURE 6.16**).

FIGURE 6.12

High School Level Implementation of the Classroom Assessment Cycle: Tenth Grade, Mrs. Green's Class

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
UNIT GOALS: "As a result of activities in grades 9–12, all students should develop understanding of ❑ Personal and community health ❑ Population growth ❑ Natural resources ❑ Environmental quality ❑ Natural and human-induced hazards ❑ Science and technology in local, national, and global challenges" (National Research Council, 1996, p. 193) In addition, under the guide to "abilities necessary to do scientific inquiry," (Content Standard A), students will ❑ "Recognize and analyze alternative explanations and models ❑ Communicate and defend a scientific argument" (National Research Council, 1996, pp. 175–176).	Today's lesson comes at the end of a 10-day, problem-based learning experience. On the first day of this experience, students were presented with a real-life problem in the form of a memo from a hospital chief executive officer to the hospital administrator. (See **FIGURE 6.13** for a copy of this memo.) This is the first problem-based experience for the students. They are working in groups of four. Each group of four is assigned a particular role to play (the group may all be medical staff, or all accountants, etc.). **FORMATIVE ASSESSMENTS** Students were asked to write a problem statement (put the problem in their own words); construct "Know" and "Need to Know" lists; conduct research to answer "Need to Know" questions; and prepare a solution to the problem. As students completed each product, she asked them to post the products on chart paper around the room. Students read all the other products and then discussed strengths and weaknesses displayed (***peer review***; ***self-assesssment***).	Mrs. Green read the problem statements written by the students and noticed that many students simply jumped into planning how to separate the twins. They missed the part of the memo that states "consider the feasibility" of this undertaking. The problem is designed to have students consider alternatives and ethics, not to have them debate biology (anatomical structures). Therefore, re-focusing the students on the feasibility aspect was necessary for the learning targets to be addressed.	Mrs. Green had students share their problem statements. Through peer review, the problems were refined and re-focused on the "feasibility" of the undertaking, rather than on "how to" separate the twins. Mrs. Green provided sample, exemplary products (problem statements, Know and Need to Know Lists) constructed by students from past years working on a different problem (***written exemplars***).

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
TODAY'S LESSON: SPECIFIC OBJECTIVES (kid-language) Students will ❑ Make oral presentations on the conjoined twin problem (See the rubric for this presentation in **FIGURE 6.14**) ❑ Evaluate all proposed solutions and select or formulate the "best" solution. Justify reasons this would be the "best." (See the directions for this justification in **FIGURE 6.15**.) These specific objectives are aligned with the "Science and technology in local, national, and global challenges" portions of Content Standard F (above) as well as with the "abilities necessary to do scientific inquiry" one. **LEARNING TARGETS** Knowledge (scientific background knowledge is necessary for students to choose best solution) ***Reasoning*** (students evaluate, justify) ***Process Skills*** (oral presentation) ***Product*** (justification paragraph)	Mrs. Green also wrote comments or suggestions on the products (***written feedback***). Students were given the opportunity to improve their own products after this discussion.		
Mrs. Green displays today's objectives on the overhead projector screen and reviews these with students at the beginning of class.	Today, students will make oral presentations (***constructed response—performance assessment***) of their proposed solutions. Students will receive feedback on their performances via an oral presentation rubric scored by the teacher and by the other student groups. Since there are five student groups, each group will receive four rubrics scored by student groups and one rubric scored by the teacher.	Mrs. Green notes that the student groups are actually giving lower scores on the rubrics than she is giving. There is consistency (inter-rater reliability), however, between the dimensions she scored the lowest and the ones the students scored the lowest. The two dimensions earning the lowest scores are teamwork and content organization.	Mrs. Green will present her analysis to the class and let them know that teamwork and content organization generally received the lowest scores. She will ask students to brainstorm ways that these areas can be improved. Working with the students, she will create a list of strategies students can use to enhance scores in these areas.

CLARIFY LEARNING TARGETS	GATHER EVIDENCE IN A VARIETY OF WAYS	ANALYZE ASSESSMENT DATA	MODIFY INSTRUCTION
	FORMATIVE ASSESSMENTS After all groups have presented, Mrs. Green helps students make a list of all proposed solutions to the problem. She posts this list on the overhead. Mrs. Green then asks (***questioning***) students to select the ONE best solution and justify the selection of this solution. She hands out the directions for the activity (See **FIGURE 6.15**), reviews the directions with the students, and explains the rubric (***written feedback***) for the assignment. Students are given the opportunity to ask questions about the rubric. Students are given a sample of student work from a previous class, and they score this sample work (***peer review***) with the Justification rubric. Mrs. Green facilitates a class discussion (***teacher/student dialogues***) to reach consensus on the rubric scores for the exemplar. She asks (***questioning***) students to explain their scores.	After reading through the student papers for the justification assignment, Mrs. Green finds that students have difficulty articulating four pro and four con statements as well as difficulty identifying possible consequences of solution implementation.	Tomorrow, Mrs. Green will ask students to form groups of two. Each student in the group will review the other's justification assignment, score it with the rubric, and make suggestions for improving it. The suggestions must refer to the dimensions found on the justification rubric. Students will then be given an opportunity to revise their assignment. After this revision, students will hand in the assignment and Mrs. Green will score it with the rubric as well as writing comments on the papers.

FIGURE 6.13
Problem Presentation: Hospital Memo

Inter-Office Memorandum

To: Dr. Kay Emko
Hospital Administrator

From: Dr. Jim Reed
Chief Executive Office

Date: Nov. 14, 2005
Re: Smith conjoined twins

As you know, Dr. M.E. Seay has contacted this office about the possibility of performing separation surgery on conjoined twins. These twins, born in Romania, were adopted this month by Wilbur and Darlene Smith. I am forwarding to you an anatomical schematic, provided by Dr. Seay, which shows the body structures shared by the twins. Please assemble a planning team, which will include medical staff, accountants, legal experts, hospital facility (physical plant) representative, and counselors to consider the feasibility of the hospital pursuing this undertaking.

The hospital board will meet on Nov. 23, 2005, at 9:00 a.m. I would like for your team to present its recommendations at this meeting.

The attached anatomical schematic shows the following:

- The twins are two children above the waist and one below
- 1 rib cage
- 1 liver
- 2 gall bladders
- 1 large intestine
- 1 small intestine
- 1 pair of ovaries
- 1 pelvis
- 1 vagina
- 1 urethra
- 2 spinal cords that fuse at the base
- 3 or 4 lungs
- 2 hearts with a common bloodstream
- 2 stomachs
- 1 right kidney
- 2 left kidneys
- 1 uterus
- 1 bladder

FIGURE 6.14
Rubric for Conjoined Twins Oral Presentation by a Student Group

CRITERIA	EXCELLENT	ADEQUATE	POOR
DELIVERY	Relaxed, self-confident, and appropriately dressed for purpose or audience; builds trust and holds attention by direct eye contact with all parts of the audience; fluctuations in volume and inflection help to maintain audience interest and emphasize key points.	Quick recovery from minor mistakes, but some tension or indifference apparent; appropriately dressed; occasional but unsustained use of eye contact with audience; uneven volume; variation in inflection is present, but not consistently implemented.	Nervous tension obvious; inappropriately dressed; no effort to make eye contact with audience; low volume and/or monotonous tone causes audience to disengage.
INTRO-DUCTION	Effectively captures interest/attention of the audience; lays out plan and sequence of presentation; informs audience of group role.	Lays out plan and sequence of the presentation, but no attention-getting "hook" is present, informs audience of group role.	Introduction is missing—no plan for the presentation is given. Group role not identified.
CONTENT ORGANIZATION	Shares the group's problem statement, explains why the group chose this statement; proposes a solution to the problem, explains rationale for the solution, justifies why solution is sound and workable.	Shares problem statement; proposes a solution; justifies the solution.	Problem statement is missing; group proposes a solution but fails to justify why it is a sound and workable solution.
COMPREHENSION	Understood problem well enough to answer questions posed by audience; provided quality answers to question—answers were based in facts, clearly related to question asked, and provided audience with needed information.	Attempted to answer questions, but answers were of less quality—occasionally facts were misrepresented or inaccurate.	Unable to provide answers to audience questions.
TEAMWORK	Each member of team provides an overview of his/her contributions to the solution of the problem; each team member speaks during the presentation; all team members field questions at the end from the audience.	Each team member speaks during the presentation; presentation provides evidence that most team members contributed to the solution.	Not all team members speak during the presentation; contributions of individual team members not addressed.
QUALITY OF SOLUTION	Solution aligns with problem statement (actually solves the problem presented); solution addresses all constraints listed in original problem statement; solution is authentic to group's role; solution honors deadline and budget; solution is feasible within the time frame given in the problem and is completely supported by research.	Solution aligns with problem statement; solution addresses most of the constraints; solution is authentic to group's role; solution honors deadline and budget; solution is partially supported by research.	Solution fails to align with problem statement; fails to address majority of constraints; solution is not authentic to group's role; solution fails to honor either the given deadline or budget; research support for the solution is missing.

FIGURE 6.15
Directions for Justification

1) From all the solutions proposed today, select the two "best" solutions. Briefly describe these solutions.
2) Construct a "pro" and "con" chart for each of these two proposed solutions.
3) List positive and negative consequences of implementing each solution.
4) Write a paragraph, stating the one "best" solution and explain your rationale for choosing this one over the other one.
5) Use the following rubric to maximize your performance on this justification task.

CRITERIA	GOOD	ADEQUATE	POOR
SOLUTION SELECTION	Solutions are selected from those presented today; solution descriptions are accurate, complete, and clear.	Solutions are selected from those presented today, solution descriptions are accurate, but may be incomplete or lacking in clarity.	Solutions are selected from those presented today, but descriptions are missing or hard to understand (insufficient detail is given to check accuracy of solution description).
PRO/CON CHART	At least four "pro" and four "con" statements are listed for each of the two proposed solutions; "pros" and "cons" accurately reflect implementation issues	At least three accurate "pro" and three accurate "con" statements are listed for each of the two proposed solutions..	Less than three accurate "pro" and three accurate "con" statements are listed for each of the two proposed solutions.
CONSE-QUENCES	At least two positive and two negative outcomes (consequences) of implementing each of the two solutions is listed and explained.	At least two positive and two negative outcomes of implementing each of the two solutions is listed.	Less than two positive and two negative outcomes of implementing each of the two solutions is listed.
JUSTIFICATION PARAGRAPH	Paragraph clearly identifies and explains the "best" solution; student provides at least three viable reasons for choosing this solution over the other. Paragraph contains no grammatical errors.	Paragraph clearly identifies and explains the "best" solution; student provides at least two viable reasons for choosing this solution over the other. Paragraph contains no more than two grammatical errors.	Paragraph identifies the "best" solution, but fails to explain it. Less than two viable reasons are given for choosing this solution over the other. Paragraph contains more than two grammatical errors.

FIGURE 6.16

Summary of Classroom Assessment Elements Found in Mrs. Green's Tenth-Grade Biology Class

LEARNING TARGETS	ASSESSMENTS	FEEDBACK TO STUDENTS	ASSESSMENT DATA ANALYZED	INSTRUCTIONAL MODIFICATIONS
PRODUCT ❑ Creating problem statements, "Know" and "Need to Know" lists **REASONING** ❑ Conducting research to answer "Need to Know" questions	❑ Peer and self-assessment (students examined others' products and discussed strengths and weaknesses of these and of their own products) ❑ Constructed response—product assessment (creation of the products)	Verbal feedback from peers and teacher Written feedback from teacher in the form of comments	Student products	**CONTENT DIFFERENTIATION** ❑ Provide concrete examples (sample products from previous classes) **PROCESS DIFFERENTIATION** ❑ Peer review of products (problem statements, "Know" and "Need to Know" lists)
KNOWLEDGE, REASONING, PROCESS SKILLS, PRODUCT ❑ Compiling scientific background knowledge ❑ Using background knowledge to evaluate proposed solutions to the problem ❑ Justifying choice of "best" solution ❑ Making oral presentation ❑ Creating justification paragraph	❑ Constructed response—performance assessment (oral presentations) ❑ Questioning (Select one best solution) ❑ Teacher/Student dialogues (class discussions) ❑ Constructed response—product assessment (justification assignment)	Written feedback via rubric scored by peers and teacher Verbal feedback from teacher, written feedback via rubric Verbal feedback from teacher Written feedback via rubric and teacher comments	Rubric scores; student products	**CONTENT DIFFERENTIATION** ❑ Breaks down content into small "chunks" by listing all proposed solutions **PROCESS DIFFERENTIATION** ❑ Peer review of products and performances (oral presentation, justification assignment) ❑ Ask students to make and justify judgments (justification assignment) **PRODUCT DIFFERENTIATION** ❑ Students express knowledge in a variety of ways (discussions, oral presentations, justification assignment)

The examples in Chapter 6 provided insights into ways teachers implement the Classroom Assessment Cycle to enhance student performance in science. The focus was on formative assessment (assessments that provide feedback to students while they are learning.) Therefore, the classroom snapshots did not address grading issues. For an overview of productive grading practices, you may wish to read the Appendix below.

Appendix: Grading and Reporting Student Performance in Science

In Chapter Five, we examined methods to use for analyzing assessment data and then various means for modifying instruction based upon such analyses. In this chapter, we will concentrate on ways to provide meaningful performance reports to both students and parents. In order to offer such meaningful reports, we must first review the three purposes of assessment: Diagnostic, Formative, and Summative.

These three terms were first introduced in Chapter Four. The definitions for the terms were:

> ***Diagnostic Assessments***: Used to ascertain the current level of understanding that students possess about a concept prior to instruction.
>
> ***Formative Assessments:*** Used to provide feedback to students as they progress toward a learning goal, during the learning process.
>
> ***Summative Assessments:*** Used to rate the proficiency of students relative to a particular skill or skill set at the end of instruction.

Diagnostic assessments are primarily used by teachers to aid in planning meaningful instruction. While the results of such assessments may be shared with students or parents, it is usually the case that teachers alone review these results. Students, however, are the primary audience for reports of formative assessments, information that is often communicated to them through the use of verbal feedback, written feedback or rubrics. Through such feedback mechanisms, students come to understand their strengths and weaknesses. Of course, the results of such formative assessments may also be shared with parents. Summative assessment reports, those coming at the end of a grading period, are usually more formal than formative ones, and they are distributed in a format usually chosen by the district or county educational agency. This agency may also prescribe the frequency of such reporting (as every three, six, and nine weeks, etc.). Summative grades are those grades that we find on students' report cards.

In this appendix, we shift our focus from formative feedback or "scoring" to "grading" (summative feedback). We will consider how to enhance the reliability of summative grades found on student report cards, and we will examine methods of calculating such grades. Finally, we will look at a grade book format that supports productive grading plans.

Grade Reliability

Grades are reliable when they give an accurate picture of student achievement relative to a learning target or standard. Brookhart (2004, p.10) describes this as:

> The grade students get for an assignment should be the closest possible estimate you can make about their real achievement on whatever the assignment is supposed to measure. It shouldn't be a fluke—a low score underestimating achievement because a student was ill, or a high score overestimating achievement because the student made a lucky guess studying only the one question you happened to pick for a test. It shouldn't be an accident of which teacher grades the paper—a "hard grader" or an "easy grader." The grade should indicate the level at which students would generally register on any assignments designed to indicate achievement of the same learning goals.

In this quote, Brookhart emphasizes two important components of grades: accuracy and consistency. Like other assessment experts, (O'Connor, 1999 & 2002; Marzano, 2000; Stiggins, 1993; Gusky, 2004), Brookhart insists that grades should be a measure of student achievement relative to some standard. If we accept this grading definition, then grades should not include such non-academic factors as attendance, conduct, cooperation, ability, or attitude. O'Connor (1999) suggests that such non-academic factors be recorded and reported separately.

In addition to basing summative grades on academic performance, four other factors highly impact grading accuracy and consistency. These are:

1) Which assignments are "counted."
2) How much these assignments count (weighting).
3) The statistical method used to calculate final grades.
4) The use of zeros.

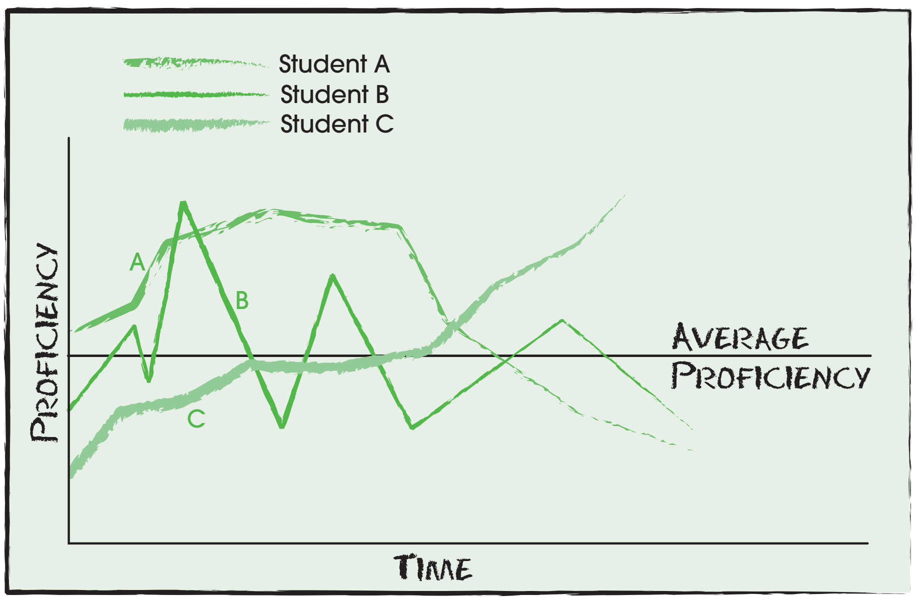
What Counts?

To decide which grades should "count," (that is, which assignment grades will be used to create the summative, report card grade), we must first return to the three purposes of assessment: diagnostic, formative, and summative. Should grades on diagnostic assignments be used to figure final grades? Most teachers would answer "no" to this question. Diagnostic assignments are given prior to any instruction. They are simply the measure of what a student already knows about a concept or subject. The purpose of diagnostic assessments is to plan appropriate teaching. We use diagnostics to prevent overdwelling on concepts that students already understand and to focus our instruction on possible misconceptions and/or perceived weaknesses in student understanding. Diagnostic assessment results are not usually used in calculating summative grades.

The use of formative scores is more problematic, however. Even though, by definition, formative scores are taken while students are still *forming* their understandings of concepts, teachers routinely record these scores as grades and then use them to calculate overall, final grades for the report card. Whether this results in accurate, consistent grades generally depends on where a particular student is located on the learning curve when the grade is assigned. If the student has had time to understand the concept presented and has been afforded sufficient practice time, then the grade may accurately reflect student achievement. The reverse may be true, however, for students lower on the learning curve!

Ken O'Connor dramatically illustrates the problematic nature of using formative assessments in the chart below. (See **FIGURE A.1**.) In this graphic, we find the proficiency levels of students learning to pack parachutes measured over time. Student A begins with above-average proficiency, gains in proficiency, and then decreases in proficiency at the end of the recorded time period. Student B shows a lack of consistency in performance, with proficiency rising and falling many times throughout the recording period. Student C began with very low proficiency, but has gradually increased in proficiency throughout the recording period. The question arises: At the end of the recording period, whom do you want to pack YOUR parachute?

FIGURE A.1
Parachute Packers



O'Connor, K. (2002). How to grade for learning. *Thousand Oaks, CA: Corwin Press. p.33, Figure Intro.11*

When asked this question, most participants in O'Connor's workshops will answer that they would like Student C to pack their parachute. Student C has gradually improved in performance and has the highest proficiency rating on the day of the jump. O'Connor then asks teachers to speculate about the students' summative grades. Which student would have the highest grade point average at the end of the recording period? Most participants agree that Student A would have a higher average, since Student A has more data appearing over the "average proficiency" line on the graph.

This "Parachute Packing" graphic, then, illustrates the problematic nature of using formative grades to calculate ending grades. Should Student C have a lower average than Student A, simply because he took longer to learn? Should all the data points on the graph (some taken while students were still learning, i.e., formative grades) be used to figure the final, overall grade? O'Connor emphasizes that if we simply want to know if the students can pack a parachute, the use of all data points is unnecessary and actually unfair, particularly to Student C. O'Connor further stipulates, "I believe formative assessment has no place in determining report card grades" (O'Connor, personal communication, October 3, 2005).

How Much Does the Assignment Count (Weighting)?

It is not uncommon for teachers to weight assignments differently. By "weighting," teachers assign relative importance to assignments. For example, a teacher may count homework as only 10% of the final grade, but the final exam may count 50%. Such differential weighting can have a huge impact on the final grade. Let's look at a research study conducted by Marzano (2000).

In this study, Marzano asked two teachers who were team-teaching the same set of students to figure final grades for their 26 students. Both teachers used the same set of recorded grades and both worked independently to calculate ending grades. When the grades for the 26 students were compared, Marzano (2000, p. 5) found:

- One student was assigned grades that differed by three letter grades (e.g., A vs. D).
- Two students received grades that differed by two letter grades (e.g., A vs. C).
- Eight students' grades differed by one letter grade (e.g., A vs. B)
- Fifteen students received the same grade from both teachers.
- In all, there was agreement on only 15 of 26 grades or 57.7% of the grades.

Marzano found that the differences in grading arose from the weighting of the assignments. One perhaps counted performance assessments as more important than homework assignments, or one weighted the chapter test more than the other.

This research study illustrates how weighting of assignments can impact summative grades. The same data set of assignments can result in widely divergent ending grades, depending on the weighting scheme used. To ensure valid grades, the weighting scheme should be one that supports accuracy and consistency and one that is closely aligned with academic standards.

Statistical Methods Used

Throughout recent grading history, teachers have tended to average all grades in order to get final grades. "Averaging" uses the statistical method of finding the *mean*. However, there are two other statistical methods that could be used in calculating final grades. These are the mode and the median. Let's take a set of scores for one student and compare the final grade obtained from using all three statistical methods. Here are Taquisha's scores for the third grading period in science: 93, 94, 93, 67, 98, 91, 90. Using the mean (or "averaging") to figure Taquisha's grade, we would add them all together and divide by seven (the number of grades). By doing this, we would obtain an average of 89.4 for Taquisha. On a typical grading scale of 80–89% = B, Taquisha would have a very high B.

Using the statistical method of the *mode* means that we simply look for the grade that occurs most often. For Taquisha, this grade would be 93, since she earned this grade twice. Therefore, if we use the mode to figure Taquisha's final grade, she would get a 93 (low A) on her report card.

Using the *median* is the third choice we have. To find the median, we must first put Taquisha's grades in numerical order, from highest to lowest. This gives us: 98, 94, 93, 93, 91, 90, 67. The second step is to look for the grade "in the middle." Since there are 7 grades, we look for the grade that would have 3 grades above it and 3 below. (If Taquisha had had an even number of grades, we would simply average the two in the middle to find the median.) This "middle" grade is 93. Therefore, if we use the median to calculate Taquisha's final grade, she would again receive a 93 (low A).

Which of these three statistical methods leads to the most meaningful (or appropriate) summary of Taquisha's work? If we look at all of Taquisha's grades, on a grading scale where an A = 90–100%, Taquisha only has one grade that is not in the A range (67). This one grade of 67 results in dropping Taquisha's final grade from an A to a B when using strict averaging (the mean) to figure Taquisha's grade. The other two statistical methods both calculate A's for Taquisha.

Teachers must use their professional judgment to decide how to calculate final grades. Informing this judgment should be the question, "Which method results in a grade that most closely aligns with the student's actual academic performance?" If an outlying score (one significantly different from most of the others) appears to unduly influence the final grade, the teacher may wish to examine other statistical means of calculating final grades. Teachers may also want to use caution in implementing the mode as a means of figuring the final grade. The mode is fairly unstable as an indicator of student achievement. Consider the following set of grades for Kenon: 92, 67, 94, 96, 99, 67, 91. Using the mode in Kenon's case would result in a grade of 67, as opposed to a mean of 86.6 or a median of 92. Actually, Kenon's two grades of 67 appear to be outliers (unusual grades for Kenon). Therefore, the strict use of the mode can result in grades that do not accurately reflect student achievement relative to standards.

The Use of Zeros

Our working definition of grade reliability states that grades accurately depict student academic performance relative to a learning target or standard. A grade of zero rarely provides such an accurate reflection of what a student has learned. Zeros are most commonly used to denote absent or missing assignments. However, when such zero scores are used to calculate summative, report card grades, their effects may be greatly magnified and therefore, zeros may negatively impact grade reliability. Let's look at a set of scores for Jason.

In science class this term, Jason has earned the following set of marks: 90, 92, 88, 90, 91, 89, and 91. Simple averaging of these scores gives Jason an overall final grade of 90.1. Adding in a single zero to this set of scores, however, drops this average to 78.9. Which final score more accurately reflects Jason's performance in science this term?

Teachers often defend the use of zeros by stating that they cannot give students credit for missing work. While this is true, many fail to assess the huge negative impact that even one zero can have on a student's ending average. How can the effect of zero scores be minimized?

O'Connor (1999) states that teachers consider giving the highest possible "F" score on a missing assignment, rather than recording a zero. For example, if the grading scale uses "0–69" as an "F," the teacher would record "69" for all missing work. Conversely, Guskey (2004) suggests that teachers consider giving "Incomplete" grades until students turn in missing assignments. Students with Incomplete grades could then be required to attend special after school or Saturday sessions. Another of Guskey's (2004, p. 52) suggestions is to change the grading scale used. He states:

> Schools using this approach shift from percentage grading scales where, for example, A = 90–100%, B = 80–89%, C = 70–79% and so on to whole number scales where A = 4, B = 3, C = 2 and so on. In other words, although teachers can still assign zeros to student work that is missed, neglected, or turned in late, the effect of a zero is lessened because it is not so extreme.

Using Guskey's change in grading scale suggestion, let's return to Jason's grades:

90 = A = 4
92 = A = 4
88 = B = 3
90 = A = 4
91 = A = 4
89 = B = 3
91 = A = 4
0 = F = 0

On this scale, Jason's average would be 3.25, corresponding to a B average. This B average more closely reflects his overall work in science this term than the 78.9 (C) he would receive on the percentage grading scale, using the mean (including the zero) to calculate the final grade.

We can also use O'Connor's suggestion and replace Jason's zero with 69. Doing this gives Jason an ending average (using grades of 90, 92, 88, 90, 91, 89, 91, and 69) of 87.5 (B). Again, this grade appears more consistent with his performance in science than the final grade of 78.9. We should note, however, that this suggestion by O'Connor is hugely controversial. Some districts who have implemented this suggestion have found grades challenged by school boards and by parents.

RESOURCES

Reeves, D.B. (2004). The case against the zero. *Phi Delta Kappan*, *86*(4), 324–325.

ZAP Schools (Zeros Aren't Permitted):

- Kruger School of Applied Technologies (middle school), San Antonio, TX
- Tri-Cities High School, Fulton County, GA

Let's return to Jason's grades for a moment and re-examine the use of different statistical methods to calculate a final grade. Jason has one zero. His other grades are: 90, 92, 88, 90, 91, 89, and 91. We've already established that using the mean (averaging) will result in a final grade of 78.9. What if we use the median? Placing Jason's grades in numerical order gives us: 0, 88, 89, 90, 90, 91, 91, and 92. The median is 90.

In summary, then, teachers can minimize the impact of zeros using several different strategies. They can: 1) use the highest possible "F" score in place of the zero; 2) give "incomplete" grades for missing work; 3) use a whole number scale rather than a percentage scale; or 4) use the median to figure the final grade, rather than the mean. Teachers should use their professional judgment to decide which method will work the best in their classroom and which will best align with school, district, or state policy.

The above discussion of grade reliability has brought to light some problematic areas in grading. We have seen some very familiar and often used grading practices that may negatively impact grade reliability. Rather than dwelling on problems, however, this publication attempts to propose solutions. The following summary from Butler & McMunn (2006) outlines a productive grading plan that will result in reliable grades.

FIGURE A.2
Creating a Productive Grading Plan

IDEA	EXAMPLE/EXPLANATION
1. Grades are strictly tied to student achievement of curricular standard.	Grades contain only information about student academic performance—no data on conduct, attendance, or attitude are included.
2. Grades are meaningful and derived from quality assessment tasks.	Valid assessments aligned with the standards are used to collect data. Example: A performance assessment is used to determine if a student can correctly operate a microscope (not a pencil-and-paper test). Sufficient data (multiple sources, triangulation) exists.
3. Grades are figured from summative assessments.	Assessment data is collected while students are learning new concepts, but such formative data are not used in figuring final grades. Grades on assignments are only taken (and counted) after students have had time to practice the new knowledge or skill and have received feedback on their performances.
4. Zeros for missing work do not overly penalize students.	Either O'Connor's or Guskey's suggestions for minimizing zeros are used.
5. Grading policies are developed and shared with students and parents.	Written grading policy statements are distributed to parents and students at the beginning of the year. Both parents and students are given opportunities to ask questions about the policies in order to clarify their understanding.
6. Grading policies are standardized throughout a school/district.	Schools and/or districts have regularly scheduled meetings to discuss grading practices. Teachers who teach the same subjects at the same grade levels have opportunities to work together and discuss scoring of sample student work.
7. Grades reflect students' current achievement levels.	Use the median as the statistical method by which to calculate final grades ***or*** weight the most recent work more than earlier work ***and/or*** only use summative assessments for figuring final grades.

Grade Book Formats

In order to implement such a productive grading plan as the one outlined in **FIGURE A.2**, changes in record-keeping may be necessary. **FIGURE A.3** provides one sample grade book format that may be useful in visualizing new ways of record-keeping. Note that this grade book is divided into two sides: A Formative Page and a Summative Page. We'll examine each page in detail.

FIGURE A.3

Sample Grade Book Format for Six-Week Grading Period

Topic Covered: Weather

Formative Page

Students	Diagnostic Journal Entry on Weather Concepts	Taking Measurements: Thermometer	Taking Measurements: Anemometer	Taking Measurements: Barometer	Classifying Cloud Types	Weather Map—Symbol Interpretation	Predicting and Forecasting	Concept Map	Terms Worksheet
1	√+	a	nc	a	8/10	7/7	4	4	18/20
2	√	nc	nc	I	9/10	6/7	4	3	Missing
3	√-	I	a	nc	10/10	Missing	3	4	17/20
4	√-	I	I	I	7/10	5/7	2	3	14/20
5	√	a	a	a	8/10	4/7	3	2	20/20

Summative Page

Students	Performance Assessment: Accurate Measurements (Lab Practical Exam)	Terms Quiz (Matching)	Clouds Quiz (Label)	Forecasting Essay	Multiple-Choice Test	Scores converted to 4-point scale	Weighted Scores = (Score converted to 4-point scale) x (weight percentage)	Final Weighted Score	Final Grade
	(30%)	(10%)	(10%)	(30%)	(20%)				
1	98	100	100	24=A	95	4,4,4,4,4	1.2, 0.4, 0.4, 1.2, 0.8	4	A
2	96	94	84	22=B	86	4,4,3,3,3	1.2, 0.4, 0.3, 0.9, 0.6	3.4	B
3	78	96	88	19=C	92	2,4,3,2,4	0.6, 0.4, 0.3, 0.6, 0.8	2.7 (rounds to 3)	B
4	70	96	98	18=C	78	2,4,4,2,2	0.3, 0.4, 0.4, 0.6, 0.4	2.1	C
5	100	100	100	24=A	96	4,4,4,4,4	1.2, 0.4, 0.4, 1.2, 0.8	4	A

Formative Page

On the Formative side of the grade book, we find several different entries. The first is diagnostic, rather than formative and consists of a journal entry on weather concepts. The teacher has graded this with a ✓+, ✓, ✓- grading scheme. The teacher is not trying to record a grade for the diagnostic, but rather record a grade that will help her remember the relative knowledge levels shown by the students. We assume that students making a ✓+ would have shown greater knowledge or understanding of weather concepts on this diagnostic than those making ✓-. We hope that the teacher analyzed the answers from all the students to help her plan instruction on this unit.

The next few entries on the Formative side of the grade book appear to be performance items, in which students were required to take accurate measurements with various weather instruments. For these measurements, the teacher has recorded "A," "NC," or "I" grades. "A" stands for Accurate measurement, "NC" is for Inconsistent accuracy, and "I" is for Inaccurate measuring. Hopefully, students scoring NC or I were given further opportunities to practice and meaningful feedback to help them improve.

The next two grades on the Formative Side (Classifying Clouds and Weather Map-Symbol Interpretation) are simply recorded as "number right/total number." A similar grading scheme was used on the last assignment recorded, the Terms Worksheet. Again, this information is for the purpose of determining who may need further practice in these areas. Note that two different students are missing assignments in these categories.

The teacher appears to have used a rubric for the "Predicting and Forecasting" and "Concept Map" grades. These rubrics divided students into proficiency levels of Excellent (4), Good (3), Adequate (2), and Re-do (1). We note that no students were required to re-do the assignments. The rubrics were analytical, with clear descriptions of each level of performance. Therefore, when students received the rubric scores, they could clearly identify their own strengths and weaknesses. The written descriptors could also aid the students in improving their performances.

We should note that none of the assignments listed on the Formative side of the grade book are used to calculate the final grade for the six-weeks grading period. The Formative side of the grade book is strictly informational in nature. It helps the teacher capture snapshots of student achievement during the learning process. He can use such snapshots to differentiate instruction for his class.

Summative Page

The summative side of the grade book contains the grades that "counted." In other words, these are the grades that were used to determine the final six-weeks grade. On this side, we find five summative grades, each showing the weight of the assignment. There is a lab practical exam that counts 30% of the final grade, two quizzes (Terms and Clouds) that each count 10% of the final grade, a Forecasting Essay counting 30% of the final grade, and a Multiple-Choice Test counting 20% of the final grade.

Scores for all assignments except the Forecasting Essay are recorded as percentage grades. A rubric was used to grade the Forecasting Essay, and the grade conversions are given where the number of points earned were translated into letter grades.

All grades were then converted to a four-point scale, with A = 4, B = 3, C = 2, D = 1, and F = 0. Finally, weighted scores were calculated into the final weighted score, using the formula indicated, and a final grade was assigned.

How does this sample grade book format support the productive grading plan outlined in the previous section? Let's look at this point by point:

1. **Grades are strictly tied to student achievement of curricular standard.**

   The curricular standards being taught by this teacher come from the National Science Education Standards for K–4. The first is under the subtitle of Earth and Space Science (National Research Council, 1999, p. 134). One of these content standards reads, "Weather changes from day to day and over the seasons. Weather can be described by measurable quantities, such as temperature, wind direction and speed, and precipitation." Two other content standards from NSES are also being utilized by this teacher. These fall under the Science as Inquiry category and in the Abilities Necessary to Do Scientific Inquiry subcategory (National Research Council, 1999, p. 122). These are stated as: "Employ simple equipment and tools to gather data and extend the senses," and "Use data to construct a reasonable explanation."

   The assignments given to the students in this six-week period align with one of these national standards.

2. **Grades are meaningful and derived from quality assessment tasks.**

   Without viewing the actual sample of assessments used, this is harder to discern. However, we are able to tell that the teacher provides opportunities for practice before taking summative grades and that the formative assessments do align with the summative ones. The teacher utilizes rubrics to provide meaningful feedback to students.

3. **Grades are figured from summative assessments.**

   No formative grades have been used to figure final grades. All grades that "count" are summative ones.

4. **Zeros for missing work do not overly penalize students.**

   No zeros were recorded in the summative section of this grade book. However, even if there were recorded zeros, the effect of the zero grades would be minimized as the teacher converts all grades to a four-point scale before figuring final grades.

5. **Grading policies are shared with students and parents.**

   We have no data from this grade book to tell us if the teacher shared grading policies with students and parents.

6. **Grading policies are standardized throughout a school/district.**

   We have no data that either supports or refutes this action by the teacher.

7. **Grades reflect students' current achievement levels.**

   The teacher uses the median as the statistical method for figuring final grades and only uses summative assignments to compute final grades.

By examining each productive grading practice, this analysis of the sample grade book finds that five of the seven recommended grading practices are exemplified by this teacher.

APPLICATION

Examine your own grade book. Does it exemplify a productive grading plan? If not, what changes could you make?

REFERENCES

Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman, Inc.

Black, P. & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80* (2), p. 139–148.

Black P., Harrison, Christine, Lee, Clare, Marshall, Bethan, and Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, (9), p. 9–21.

Bloom, B.S. (1956). *Taxonomy of educational objectives, the classification of educational goals, cognitive and affective domain*. New York: David McKay Co.

Borich, G. (2000). *Effective teaching methods*. Upper Saddle River, NJ: Prentice Hall.

Brookhart, S.M. (2004). *Grading*. Upper Saddle River, NJ: Pearson Education, Inc.

Brown, G. & Edmondson, R. (1984). Asking questions. In E. Wragg (Ed.) *Classroom teaching skills* (pp. 97–119). New York: Nichols.

Butler, S.M. & McMunn, N.D. (2006). *A teacher's guide to classroom assessment: Understanding and using assessment to improve student learning*. San Francisco, CA: Jossey-Bass, Inc.

Chappuis, S., Stiggins, R.J., Arter, J., & Chappuis, J. (2004). *Assessment for learning: An action guide for school leaders*. Portland, OR: Assessment Training Institute.

Guskey, T.R. (2004). O Alternatives. *Principal Leadership, 5* (2), 49–53.

Hibbard, K. M. (2000). *Performance-based learning and assessment in middle school science*. Larchmont, NY: Eye on Education.

Marzano, R.J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mississippi State Curriculum Frameworks. *Science frameworks* (*5th*). Retrieved August 2, 2005 from **http://marcopolo.mde.k12.ms.us/frameworks/science/sci_05.html**

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

North Carolina Department of Public Instruction. *Science standard course of study and grade level competencies*. Retrieved May 26, 2005 from **www.dpi.state.nc.us/curriculum/science/standard/toc.html**

O'Connor, K. (1999). *The mindful school. How to grade for learning*. Arlington Heights, IL: SkyLight Training and Publishing Inc.

O'Connor, K. (2002). *How to grade for learning: Linking grades to standards*. (*2nd Ed*). Arlington Heights, IL: Skylight Training and Publishing, Inc.

Regional Educational Laboratories. (1998). *Improving classroom assessment: A toolkit for professional developers*. Portland, OR: Northwest Regional Educational Laboratory.

Stiggins, R. (1993). Teacher training in assessment: Overcoming the neglect. In Wise, S.L. (Ed.) *Teacher training in measurement and assessment skills*, 27–40. Lincoln, NE: Buros Institute of Mental Measurements.

Stiggins, R. (2001). *Student-involved classroom assessment*. Upper Saddle River, NJ: Prentice-Hall, Inc.

South Carolina Department of Education. *Science curriculum standards*. Retrieved May 27, 2005 from **www.myscschools.com/offices/cso/standards/science/Sunshine State Standards**. K–2 science curriculum. Retrieved May 26, 2005 from **www.firn.edu/doe/curric/prek12/pdf/prk-2g.pdf**

Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: ASCD.

Tomlinson, C.A. & Allan, S.D. (2000). *Leadership for differentiating schools and classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.

ABOUT SERVE

The SERVE Center at UNCG, under the leadership of Dr. Ludwig David van Broekhuizen, is an education organization with the mission to promote and support the continuous improvement of educational opportunities for all learners in the Southeast. The organization's commitment to continuous improvement is manifest in an applied research-to-practice model that drives all of its work. Building on research, professional wisdom, and craft knowledge, SERVE staff members develop tools, processes, and interventions designed to assist practitioners and policymakers with their work. SERVE's ultimate goal is to raise the level of student achievement in the region. Evaluation of the impact of these activities combined with input from stakeholders expands SERVE's knowledge base and informs future research.

This rigorous and practical approach to research and development is supported by an experienced staff strategically located throughout the region. This staff is highly skilled in providing needs assessment services, conducting applied research in schools, and developing processes, products, and programs that support educational improvement and increase student achievement. In the last three years, in addition to its basic research and development work with over 170 southeastern schools, SERVE staff provided technical assistance and training to more than 18,000 teachers and administrators across the region.

The SERVE Center is governed by a board of directors that includes the governors, chief state school officers, educators, legislators, and private sector leaders from Alabama, Florida, Georgia, Mississippi, North Carolina, and South Carolina.

SERVE's operational core is the Regional Educational Laboratory. Funded by the U.S. Department of Education's Institute of Education Sciences, the Regional Educational Laboratory for the Southeast is one of ten Laboratories providing research-based information and services to all 50 states and territories. These Laboratories form a nationwide education knowledge network, building a bank of information and resources shared and disseminated nationally and regionally to improve student achievement. SERVE's National Leadership Area, Expanded Learning Opportunities, focuses on improving student outcomes through the use of exemplary pre-K and extended-day programs.

SERVE Center
at the University of North Carolina at Greensboro

P.O. Box 5367 • Greensboro, NC 27435
800-755-3277 • 336-315-7400 • Fax 336-315-7457

Ludwig David van Broekhuizen, Ph.D., Executive Director

WWW.SERVE.ORG