

**Science
Educator's
Guide to**
Laboratory
Assessment

By

Rodney Doran

Fred Chan

Pinchas Tamir

Carol Lenhardt

NSTApress®
NATIONAL SCIENCE TEACHERS ASSOCIATION





Claire Reinburg, Director
Judy Cusick, Associate Editor
Carol Duval, Associate Editor
Betty Smith, Associate Editor

ART AND DESIGN Linda Olliver, Director

Cover image, Photodisc

NSTA WEB Tim Weber, Webmaster

PERIODICALS PUBLISHING Shelley Carey, Director

PRINTING AND PRODUCTION Catherine Lorrain-Hale, Director

PUBLICATIONS OPERATIONS Erin Miller, Manager

sciLINKS Tyson Brown, Manager

NATIONAL SCIENCE TEACHERS ASSOCIATION

Gerald F. Wheeler, Executive Director

David Beacom, Publisher

Science Educator's Guide to Laboratory Assessment

NSTA Stock Number: PB145X2

ISBN: 0-87355-210-5

Library of Congress Catalog Card Number: 98-84914

Printed in the USA by Victor Graphics

Printed on recycled paper

Copyright © 2002 by the National Science Teachers Association.

Permission is granted in advance for reproduction for purpose of classroom or workshop instruction. To request permission for other uses, send specific requests to:

NSTA PRESS

1840 Wilson Boulevard

Arlington, Virginia 22201-3000

www.nsta.org

NSTA is committed to publishing quality materials that promote the best in inquiry-based science education. However, conditions of actual use may vary and the safety procedures and practices described in this book are intended to serve only as a guide. Additional precautionary measures may be required. NSTA and the author(s) do not warrant or represent that the procedures and practices in this book meet any safety code or standard or federal, state, or local regulations. NSTA and the author(s) disclaim any liability for personal injury or damage to property arising out of or relating to the use of this book including any of the recommendations, instructions, or materials contained therein.

About the Authors

Rodney Doran has been a professor of science education at the State University of New York at Buffalo's Graduate School of Education since 1969. He previously taught physics, Earth science, and math at a Minnesota high school. He holds degrees from the University of Minnesota (BS, 1961), Cornell University (MST, 1966), and the University of Wisconsin (PhD, 1969). Dr. Doran was U.S. Associate Coordinator for the Second International Science Study, and received the State University of New York Chancellor's Award for Excellence in Teaching in 1998. He has published articles on assessment in *Science and Children*, *Science Scope*, and *The Science Teacher*, and was the author of the National Science Teachers Association 1981 book, *Basic Measurement and Evaluation in Science Instruction*.

Fred Chan is a vice principal with the Toronto, Ontario, District School Board. He previously taught physics, chemistry, and biology in Guyana and the Caribbean. He holds degrees from McMaster University (BS, 1973), the University of Toronto (BEd, 1975), Niagara University (MS in Ed, 1984), and the University of Bridgeport (MS, 1989), and he expects to receive his PhD from the University at Buffalo. He is a 1996 recipient of the Prime Minister of Canada's Award for Excellence in

Science Teaching. He has written articles on assessment for *The Science Teacher* and serves on the manuscript review panel for that journal. He has also served on the National Science Teachers Association's Committee on International Science Education.

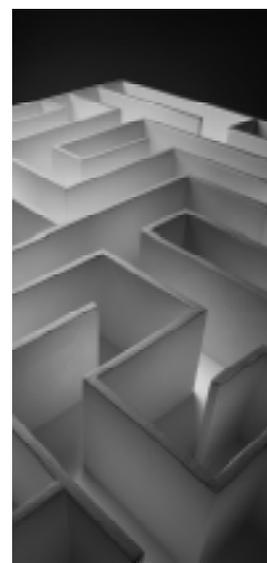
Pinchas Tamir is professor emeritus at the School of Education at Hebrew University of Jerusalem in Israel. He was also director of the Israel Science Teacher Center, where he helped introduce BSCS biology to Israel and develop the Israel Matriculation Exam for high school biology. He was Israel's research coordinator for the Second International Science Study and the Third International Science Study, and he is a visiting scholar at dozens of universities around the world. He holds degrees from Hebrew University of Jerusalem (MS, 1966) and Cornell University (PhD, 1968).

Carol Lenhardt is a retired middle school science teacher. She also was an adjunct faculty member at Buffalo State College and the University of Buffalo. She holds degrees from the University of Rochester (BS, 1956), Syracuse University (MS, 1960), and the University of Buffalo (PhD, 1994). She is a past president and fellow of the Science Teachers Association of New York State (STANYS) and coauthor of *Alternative Assessment in Science, Grade 8*.

Contents

ABOUT THE AUTHORS	v
PREFACE	viii
Background: The Importance of Assessment	viii
The Meaning of Inquiry	ix
Diagnostic, Formative, and Summative Assessment	x
High-Stakes Tests	xi
Professional Development	xi
Organization and Use of This Book	xii
ACKNOWLEDGMENTS	xv
CHAPTER 1: A Rationale for Assessment	2
The Present State of Assessment	2
The Constructivist Paradigm	3
Assessment’s Changing Nature	4
The Multifaceted Assessment System	6
Using Assessment Results—The New Paradigm	8
CHAPTER 2: Developing New Assessments	12
Toward Effective Assessment	12
An Assessment Development Model	12
State the Purpose	13
Select the Appropriate Task Format	14
Write the Task	15
Modify an Existing Task	17
Use Clear Directions and Questions	19
Consider Equity	20
Clarify Administrative Procedures	20
Develop the Scoring Rubric	22
Trial Test the Task	23
Analyze Results	23
Revise Tasks	24
Structure	24
Sequence	25
Novelty	26
CHAPTER 3: Alternative Assessment Formats	28
What Is “Alternative”?	28
Performance-Based Assessment Formats	28
Skills Tasks	28
Investigations	31
Extended Investigations	32
Student-Focused Assessment Formats	35
Graphic Organizers	35
Concept Maps	35
Venn Diagrams	38
Vee Diagramming, or Vee Heuristic	39

Portfolios	42
Oral Presentations and Debate	43
Interviews and Conferences	46
Lab Skills Checklists	47
Self, Pair, and Peer Evaluations	49
Technological Applications	53
Teacher-Directed Assessment Formats	53
Demonstrations	53
Group Visuals	54
CHAPTER 4: Using Performance Assessment Results	58
Assessment and Evaluation	58
Norm- and Criterion-Referenced Evaluations	58
Using Assessment Data	60
Scoring Performance Assessment	60
Developing a Scoring Team	69
Scorer Agreement	70
Reliability	71
Aggregating Assessment Data and Assigning Grades	72
Assessment Data Management	74
Using Results of Performance Assessment	74
Test Validity	76
Interpreting and Describing Results	77
Program Evaluation	79
Annual Plan	81
CHAPTER 5: Illustrative Assessment Tasks for Biology	84
Biology Skills Tasks	84
Biology Investigation Tasks	107
Biology Extended Investigation Tasks	129
CHAPTER 6: Illustrative Assessment Tasks for Chemistry	144
Chemistry Skills Tasks	145
Chemistry Investigation Tasks	164
Chemistry Extended Investigation Task	172
CHAPTER 7: Illustrative Assessment Tasks for Earth Science	176
Earth Science Skills Tasks	176
Earth Science Investigation Tasks	191
Earth Science Extended Investigation Task	199
CHAPTER 8: Illustrative Assessment Tasks for Physics	204
Physics Skills Tasks	205
Physics Investigation Tasks	236
Physics Extended Investigation Task	247
APPENDICES	
Glossary of Assessment Terminology	254
<i>National Science Education Standards</i> for Assessment	257
Complete Bibliography	258
INDEX	265



Preface

Welcome to the second, enlarged edition of *Science Educator's Guide to Laboratory Assessment*. This version contains fifteen new assessment tasks, and like the first edition, it presents multiple assessment formats, strategies, models, and templates appropriate for inquiry activities in the grades 7–12 science classroom and laboratory, as well as outdoors. These assessment formats and strategies are based on the most recent research on assessment, instruction, and learning and include many practical examples you can adapt for use in your classroom.

Background: The Importance of Assessment

As science teachers, we face a continual challenge of assessing what students know, are able to do, and value in learning science. Assessment provides insights into students' rates of progress in conceptual understanding, reinforces productive learning habits, and validates learning activities. Students need recurring, systematic, and regular feedback to understand their own strengths and capabilities in learning and to identify areas for improvement. We now are aware that increased use of formative assessment in science classrooms to modify teaching and to provide feedback to students has powerful positive effects on student learning. A well-designed assessment program, by providing regular and systematic feedback, goes a long way in helping students reflect on their learning. Hence the importance of assessment reform.

The assessment phase of the teaching-learning process is our primary way of “keeping score.” Teachers measure how well students learn new concepts and skills, administrators and policymakers measure the effectiveness of teaching strategies and educational and program policies, and parents use grades and marks to monitor their children's progress in school. Also, as a society we use data from assessments to compare our national progress in education with that of other nations.

There is a growing tension between the rich, authentic assessments that the science standards suggest and the increased use of large-scale, high-stakes testing. Science teachers need to come to grips with how much we teach “to the test,” and in so doing, how much we narrow the curriculum. We need to balance the requirements of high-stakes testing with designing assessments that provide students with varied opportunities to develop competencies in science and to demonstrate what they know and can do.

Assessment has become increasingly important during the past decade, as educators and policymakers seek reforms to our educational system in response to national and international priorities and challenges. Educators concerned with weak science achievement, low levels of science literacy, and poor international test scores have undertaken major reforms in science instruction. Increased international economic competition has reinforced the importance of excellence in science education as a fundamental priority for every nation to maintain its competitiveness. New insights into how children learn and

advances in learning theory, such as constructivism and the identification of alternative and prior conceptions of learning science, have added impetus to assessment reform. As a result, there is a call for widespread use of alternative assessments, and a shift away from textbook- and teacher-centered approaches to instruction.

These reform efforts, embodied in the *National Science Education Standards* (NRC 1996) and in reform documents such as *Project 2061: Benchmarks for Science Literacy* (AAAS 1993), call for widespread reform in science instruction and assessment. Old teaching strategies and assessment formats based on behaviorist theories, such as rote memorization and paper-and-pencil examinations, are being replaced with holistic, constructivist approaches that promote problem solving and higher-level thinking. These sophisticated assessments demand the use of a variety of teaching strategies to help students develop their ability to learn and to solve problems in “real-world” situations and contexts.

The Meaning of Inquiry

Inquiry has been and continues to be a concept near and dear to the hearts of science teachers. Bybee (2000) traces the long history of inquiry at least back to John Dewey in the early 1900s. Inquiry has been in and out of favor since then, depending on the reform efforts popular at a particular time.

One source of confusion about inquiry is that it is both a methodology of how scientists investigate natural phenomena *and* a methodology espoused for facilitating the engagement of students with materials and questions. To add to the confusion, process goals (to include in-

quiry) have been cited as “content outcomes” since the 1960s (Parker and Rubin 1966). This view is continued with the *National Science Education Standards* (NRC 1996), which uses *inquiry* in two ways: as abilities students should develop to be able to design and conduct scientific investigations and as the understandings they should gain about the nature of professional scientific inquiry.

Although inquiry is a mode of gathering information in many academic/scholarly fields, there are some unique aspects of *scientific* inquiry. In many of the science curriculum projects from the 1960s, inquiry was largely accepted as a collection of science processes (e.g., observing, measuring, predicting, hypothesizing). Currently it is viewed as one set of tools to further the development of scientific explanation. For instance, the *Learning Standards for Mathematics, Science, and Technology* (New York State Education Department 1996) identifies three key ideas of scientific inquiry:

- The central purpose of scientific inquiry is to develop explanations of natural phenomena in a continuing, creative process.
- Beyond the use of reasoning and consensus, scientific inquiry involves the testing of proposed explanations involving the use of conventional techniques and procedures and usually requiring considerable ingenuity.
- The observations made while testing proposed explanations, when analyzed using conventional and invented methods, provide new insights into phenomena.

Other educators have treated inquiry as virtually synonymous with problem solving and/or critical thinking. Although there is much overlap among these concepts, it may

be helpful to make the following distinctions: *inquiry* tends to focus on developing new information (relationships, concepts, principles); *problem solving* focuses on finding solutions to problems and is linked with technology; and *critical thinking*, also described as “rational reasoning,” can be considered to be a set of cognitive strategies that include, for example, deduction and induction.

In this volume, when we refer to *inquiry* we mean *scientific inquiry*. One of the clearest descriptions of the term is from the *National Science Education Standards*:

Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world. (NRC 1996, 23)

Diagnostic, Formative, and Summative Assessment

The current view is that every assessment consists of three interconnected elements—observation, interpolation, and cognition—that form a triangle. Each element is connected to and dependent on the others. Assessment tasks are designed around cognition or theories of learning. Student accomplishments provide observations and evidence for an interpretation of how much they know and can do (NRC 2001).

As we design assessment based on current theories of learning, it is important to clarify the meanings of *diagnostic*, *formative*, and *summative* assessment. The National Research Council’s Committee

on Classroom Assessment and the *National Science Education Standards* (NRC 2001) suggest that we ask the following questions to determine what type of assessment we are using:

- Where are we presently? (diagnostic assessment)
- How can we get there? (formative assessment)
- Have we arrived? (summative assessment)

Diagnostic assessment is the use of qualitative and quantitative data and information to determine where students are in terms of their knowledge and skills. The use of this assessment information tells students which areas they are strong in and which areas need academic intervention. This kind of assessment can be informal—for example, interviews, paper-and-pencil tests, and previous academic records. Diagnostic assessment is “low stakes” and answers the question “Where are we presently?”

Formative assessment is also “low stakes” and gives feedback to students about where they are in terms of their knowledge and skills. These assessments are informal and ongoing. The feedback to students should provide a roadmap for “How can we get there?” Using the roadmap, students try new ideas, look at problems differently, and discuss problems with peers and teachers. The roadmap takes us to our destination, which is the standards set forth by your state or school district.

Of our destination, we naturally ask, “Have we arrived?” That is where summative assessments enter the picture. These are culminating assessment tasks that occur at the end of a unit, topic, or course. They are considered “high-stakes” (more about this term below) because de-

cisions regarding further study, jobs, and academic standing are based on them. Summative assessments can be paper-and-pencil format or can be a collection of student work collected over time using a portfolio format. Summative assessments are of the highest stakes when the assessment data are used for credentialing purposes such as the awarding of a high school diploma.

The key distinction among these terms is the use and timing of the assessment data. Diagnostic and formative assessments are intended to support student learning. Summative assessment data are used to certify student accomplishments in terms of their knowledge and skills.

High-Stakes Tests

A few more comments on high-stakes tests are appropriate here. Just what are such tests (or assessments) from the point of view of a classroom science teacher? The key to answer this question is to determine the purpose of the assessment. When assessment results are used to give rewards to those students who obtain high test scores, then such assessments (tests) are “high stakes.” (An unwelcome result may be that those students who have low-test scores are denied educational opportunities.) Examples of common high-stakes tests are the SAT and the ACT. A recent trend in high-stakes testing is the use of state tests for graduation decisions, such as the awarding of high school diplomas. It is important that these tests satisfy test measurement principles of reliability, validity, and fairness (National Research Council 1999; AERA, APA, and NCME 1999) and that appropriate accommodations be made for English language learners and students with disabilities.

The classroom science teacher’s inclination can be to “teach to the test” in order to maximize students’ opportunities to obtain a high test score and prevent any sanctions against the school or the teacher’s performance. When the majority of class time is spent practicing and reviewing sections of previous tests, however, the curriculum will tend to narrow. In the context of high-stakes testing, good teachers know they can facilitate student learning in a variety of engaging ways (including through the use of the assessment tasks in this book), while familiarizing students with the item format and cognitive demands of the tests. In this way students are provided with the “opportunity to learn” in preparation for the tests.

It should be noted that high-stakes tests are subject to legal challenges when the test scores are used inappropriately. Test results should not be used for purposes for which the test was not designed. For example, the use of tests designed for program evaluation may be inappropriate for making decisions regarding student accountability. Increasingly, test results are being used for more than one purpose. Such use imposes limits on the consequential validity of the test. In addition, the use of the results of a single test as the sole criterion for a high-stakes decision is problematic (AERA, APA, and NCME 1999).

Professional Development

The authors share the belief that the ongoing professional development of teachers is a priority to bring alive the *National Science Education Standards*. We believe that teachers must be well grounded in their assessment knowledge and be able to use this knowledge in their classroom practice.

In the past, professional development has largely consisted of the one-day workshop where experts use “show and tell” methods to inform teachers about the latest teaching trend. We believe that lasting change in assessment practices will not come about using that disjointed approach. Teachers are professionals; they are active learners who know best what they want to know; and they see their professional development as continuous and ongoing. Our vision for effective assessment-focused professional development for science teachers is that its design must be consistent with appropriate learning theories for adults and must involve the professional’s construction of meaning and knowledge (Loucks-Horsley, et al. 1998). School districts and school administrators need to provide support in the form of time and opportunity for science teachers to meet and collaborate in ways to inform and improve classroom assessment practice.

We offer this book, now in its second edition, as a resource to assist science teachers in their ongoing professional development. Many of the ideas will challenge fundamental philosophical beliefs about learning and education. We hope that our colleagues will engage in collaborative discussions to advance their assessment practices. We envisage science teachers working with colleagues in their own schools, school districts, and professional organizations to gain expertise in assessment practices that work with their students.

Organization and Use of This Book

This book has two sections, followed by three appendices. Chapters 1-4 discuss assessment theory, research, and use, and

Chapters 5-8 contain model assessments grouped by science discipline. The following provides a brief description of what you will find in each of the book’s chapters.

Chapter 1 discusses the *National Science Education Standards* and recent research suggesting that instruction move from a primarily behaviorist approach toward constructivist models of learning and instruction. Chapter 2 addresses practical issues related to designing performance assessments that are aligned with the *National Standards*. Chapter 3 discusses the benefits and drawbacks of various assessment formats, ranging from short, focused tasks to extended investigations. Chapter 4 provides suggestions for using rubrics to establish reliable and consistent scoring of assessments, and for using data to improve both the overall science program and the performance of students.

Chapters 5-8 are disciplinary chapters that provide model assessment examples from biology, chemistry, Earth science, and physics. Most of these examples are complete tasks with information about measuring the skills appropriate for each task, time requirements, and preparing materials and equipment. There are also directions and answer sheets for students, a list of required materials and equipment, and scoring guidelines for evaluating student responses.

The disciplinary assessment tasks are grouped into three sections:

- **Skills Tasks:** relatively short, and focus on a few specific process skills.
- **Investigations:** focus on a wide variety of skills. They typically require one or two 40-45-minute class periods for completion. Students can plan and design an investigation, conduct an experiment, and com-

municate their findings and conclusions.

- Extended Investigations: last for several 40–45-minute class periods and can require several weeks for completion. Extended investigations are examples of curriculum-embedded assessments that align closely with instruction.

You are invited to use these assessments as is, modify them for specific instructional programs or purposes, or use them as models or templates to design entirely new and innovative assessments.

Although the book's primary focus is on assessing student achievement in the classroom and laboratory, we also include suggestions and examples on using these assessments for program evaluation. Many of the examples also include suggestions for revisions, depending on the uses of the assessment and the availability of materials and equipment.

There are three Appendices: a glossary, the *National Science Education Standards* for assessment, and a complete bibliography consisting of works cited and other relevant assessment resources—especially those that emphasize hands-on inquiry activities.

You can use this book *à la carte* by taking as much or as little as you desire to assist you with your assessments. You may first wish to reacquaint or familiarize yourself with the *National Science Education Standards*, principles of assessment design, and the rationale for new formats of assessment that interface with your evolving instructional pedagogy. Chapters 1–4 and the Appendices are appropriate for these purposes. Once you are comfortable with these concepts, go to Chapters 5–8 and examine the specific assessment ex-

amples that are relevant to the science disciplines you teach.

This book is practical in its approach to assessment reform. The assessments with their scoring rubrics have been field-tested by “real” teachers in “real” science classrooms. We hope you find the book useful as a resource as you continue to implement the assessment standards. We also hope you try the assessments with your students, and suggest you modify and revise the tasks to fit your needs. Involving your students at appropriate times in peer and self-reflection will help to embed your assessments in instructional practices.

Works Cited

- American Association for the Advancement of Science (AAAS). 1993. *Project 2061: Benchmarks in Science Literacy*. New York: Oxford University Press.
- American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement and Education (NCME). 1999. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Bybee, R. 2000. Teaching Science as Inquiry. In Minstrell, J. and Van Zee, E., eds. *Inquiring into Inquiry Learning and Teaching in Science*. Washington, DC: American Association for the Advancement of Science.
- Loucks-Horsley, S., Hewson, P. W., Love, N., and Stiles, K. E. 1998. *Designing Professional Development for Teachers of Science and Mathematics*. Thousand Oaks, CA: Corwin Press
- National Research Council (NRC). 1996. *National Science Education Standards*. Washington, DC: National Academy Press.
- . 1999. *High Stakes Testing for Tracking, Promotion and Graduation*. Board on Test-

ing and Assessment. Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

———. 2001. *Classroom Assessment and the National Science Education Standards*. Washington, DC: National Academy Press.

New York State Education Department, University of the State of New York. 1996. *Learning Standards for Mathematics, Science, and Technology*. Albany: New York State Education Department.

Parker, J. C., and Rubin, L. J. 1966. *Process as Content: Curriculum Design and the Application of Knowledge*. Chicago: Rand McNally.

U.S. Department of Education, Office of Civil Rights. 2000. The Use of Tests as Part of High-Stakes Decision-Making for Students: A Resource Guide for Educators and Policy-Makers. (available at www.ed.gov/offices/OCR)

essment. J. Pellegrino, N. Chudowsky, and R. Glaser, eds. Washington, DC: National Academy Press.

National Science Teachers Association. 1992. *Scope, Sequence, and Coordination of Secondary School Science, Volume II: Relevant Research*. Arlington, VA: National Science Teachers Association.

Suggested Readings

American Association for the Advancement of Science (AAAS). 1989. *Project 2061: Science for All Americans*. New York: Oxford University Press.

Black, P., and Wiliam, D. 1998a. Inside the Black Box: Raising Standards through Classroom Assessment. *Phi Delta Kappan* 80 (2): 139–48.

———. 1998b. Assessment and Classroom Learning. *Assessment in Education* 5 (1): 7–74.

National Education Goals Panel. 1996. *The National Education Goals Report: Executive Summary—Commonly Asked Questions About Standards and Assessment*. Washington, DC: National Education Goals Panel.

National Research Council (NRC). 2001. *Knowing What Students Know: The Science and Design of Educational Assessments*. Committee on the Foundations of As-

Acknowledgments

When a project of this scope is completed, it is important to recognize the many people who have contributed to it. Without the support and acceptance of our families, this book would never have happened. The help and encouragement of Shirley Watt Ireton, who was director of NSTA's Special Publications division at the time the first edition was published, was crucial at every stage. Her initial suggestions for the book's focus were both wise and timely, and her continued involvement during all phases was essential. Thanks also to Chris Findlay and the entire Special Publications staff for all their hard work on the first edition and to Judy Cusick, who was the NSTA Press's project manager for the second edition. Linda Olliver designed the cover and drew the illustrations for this later edition. Catherine Lorrain-Hale coordinated production and printing for, and laid out, both the first and second editions.

Two funded projects provided many of the assessment tasks in this book. The University at Buffalo/National Opinion Research Center joint project (UB/NORC) project developed prototype exams for high school science courses. Darrell Bock and Michelle Zimowski of

the University of Chicago's National Opinion Research Center developed the multiple-choice and open-ended items, while Rod Doran coordinated the laboratory tests at the University at Buffalo. Joan Boorman, Fred Chan, Nicholas Hejaily, and Diana Anderson focused on assessment tasks in separate fields. The New York Alternative Assessment in Science Project was a joint effort of the New York State Education Department and the University at Buffalo, producing a Teachers Guide and Collection of Tasks for grade 4, grade 8, Earth science, and biology. Douglas Reynolds, Robert Allers, and Susan Agruso were co-investigators. Dozens of teachers from western New York State helped revise and trial test the tasks presented here.

Audrey Champagne, Florence Juillerat, Cornelia Munroe, Mildred Barry, Burt Voss, Mary Kalen Romjue, Angelo Collins, Bill Williams, Lawrence Gilbert, Dwaine Eubanks, and Tony Bartley reviewed the manuscript as it progressed. Gouranga Saha and Jane Anzalone contributed in numerous ways. Special thanks to Lauri Di Matteo for single-handedly entering almost the entire contents of the book onto computer disks.

The Present State of Assessment

The roots of our current education system lie in the mass public school programs of the Industrial Revolution. The mechanized assembly lines and standardized processes that dominated that era found their way into education, where they remain deeply embedded today.

Over most of this century, school has been conceived as a manufacturing process in which raw materials (youngsters) are operated upon by the educational process (machinery), some for a longer period than others, and turned into finished products. Youngsters learn in lockstep or not at all (frequently not at all) in an assembly line of workers (teachers) who run the instructional machinery. A curriculum of mostly factual knowledge is poured into the products to the degree they can absorb it, using mostly expository teaching methods. The bosses (school administrators) tell the workers how to make the products under rigid work rules that give them little or no stake in the process. (Rubba, et al. 1991)

This assembly-line approach relies heavily upon behaviorist learning theory, which is based on three main concepts: that complex learning can be broken into discrete bits of information; that students learn by making associations between different kinds of perceptions and experiences; and that knowledge is an accumulation of discrete facts and basic skills.

Under behaviorist learning, knowledge is “decomposable” and can be broken into

its component parts without jeopardizing understanding or applicability. These decomposable skills can be learned separately using stimulus-response associations. In addition, students can learn knowledge out of context. In other words, if students demonstrate a skill in one context, they should be able to then demonstrate it in different contexts or situations. However, behaviorist learning theory does not address how discrete pieces of information are integrated into a coherent whole. Teachers must assume that students integrate this information elsewhere.

The behaviorist approach still plays a dominant role in schools, and results in learning that relies heavily on the memorization of factual information. In science education, the behaviorist legacy takes the form of teaching and learning that relies heavily on using textbooks as curriculum surrogates, and on having students memorize discrete bits of often unrelated science “factoids.” Assessments aligned with these approaches use formats made up primarily of multiple-choice, true/false, and short-answer questions. Students focus on identifying the “right” answer, as opposed to developing inquiry skills and conceptual understanding.

As a result, our education system has fallen behind in preparing students to cope successfully with the challenges of an increasingly complex and sophisticated world, a world where scientific and technological skills have become significant avenues to success. Students need opportunities to develop problem-solving and interpersonal skills if they are to succeed in this global yet “smaller” world, where

many diverse interest groups compete for increasingly scarce resources.

Science teachers are making these necessary “shifts” by implementing changes suggested by reform documents. We now use current findings from research in learning and research in science education to inform ourselves about exemplary practice. Our shifts are coupled with a move away from stimulus–response learning toward learning that is inquiry based and that focuses on previously learned science concepts, alternative conceptions, and conceptual change. Successful learning is context dependent, and is facilitated by interaction among peers. Our assessment reforms must be aligned with these instructional reforms.

The Constructivist Paradigm

A crucial aspect of this shift is to move toward “constructivist” paradigms in our design of science programs and assessments. The constructivist approach begins with a focus on what students already know about the world around them and on their understanding of this world. Using this as a base, educators work to help students develop methods for further educating themselves about the world. The end result is that students come away not only with scientific information but with an analytical way of thinking that they can apply to any number of situations in life.

Recent work in cognitive psychology suggests that meaningful learning occurs in context, and that some skills used in one context do not necessarily transfer to other contexts. Some cognitive skills are general and are used in a wide variety of academic and “real-world” tasks. On the other hand, other cognitive skills are context dependent, and apply to domain-spe-

cific knowledge and skills. There is an interface between the learning of cognitive skills and context. Some cognitive skills are transferable while others are domain specific (Perkins and Salomon 1989).

Constructivism underlies the *National Science Education Standards*, published by the National Research Council in 1996. The result of years of deliberations by educators, scientists, government officials, and a wide range of other participants, the *National Science Education Standards* view science as a process “in which students learn skills, such as observation, inference, and experimentation.” Through inquiry-based learning, “students develop under-

Figure 1.1: Assessing the Ability to Inquire or the Ability to do Scientific Inquiry. *National Science Education Standards*, NRC, 1996.

Identify Questions and Concepts That Guide Scientific Investigations

- formulate a testable hypothesis
- demonstrate the logical connections between the scientific concepts guiding a hypothesis and the design of the experiment

Design and Conduct Scientific Investigations

- formulate a question to investigate
- develop a preliminary plan
- choose appropriate equipment
- take appropriate safety precautions
- clarify controls and variables
- organize and display data
- use evidence, apply logic, and construct arguments for proposed explanations

Use Technology and Mathematics to Improve Investigations and Communications

- use a variety of measuring instruments and calculators in scientific investigations
- use formulas, charts, and graphs for communicating results

Formulate and Revise Scientific Explanations and Models Using Logic and Evidence

- formulate models based upon physical, conceptual, and mathematical concepts
- use logic and evidence from investigations to explain arguments

Communicate and Defend Scientific Arguments

- use accurate and effective means of communication, including writing, following procedures, expressing concepts, and summarizing data
 - use diagrams and charts to construct reasoned arguments
-

Figure 1.2: Changing Emphases. *National Science Education Standards*, NRC, 1996.

The *National Science Education Standards* envision change throughout the system. The assessment standards encompass the following changes in emphases:

Less Emphasis On	More Emphasis On
Assessing what is easily measured	Assessing what is most highly valued
Assessing discrete knowledge	Assessing rich, well-structured knowledge
Assessing scientific knowledge	Assessing scientific understanding and reasoning
Assessing to learn what students do not know	Assessing to learn what students do understand
Assessing only achievement	Assessing achievement and opportunity to learn
End of term assessments by teachers	Students engaged in ongoing assessment of their work and that of others
Development of external assessments by measurement experts alone	Teachers involved in the development of external assessments

standing of scientific concepts; an appreciation of the ‘how we know’ what we know in science; understanding of the nature of science; skills necessary to become independent inquirers about the natural world; [and] the dispositions to use the skills, abilities, and attitudes associated with science.” Figure 1.1 (page 3) provides an outline of standards for assessing a student’s ability to inquire or undertake scientific inquiry.

Figure 1.2 shows the changing emphases needed to promote inquiry-based learning. As you can see, the *National Standards* focus on giving students a much greater role in defining problems, designing experiments, and analyzing results. Through this process, students gain the same exhilaration of discovery that practicing scientists experience in their work when they plan and conduct investigations.

Assessment’s Changing Nature

As the nature of science education changes, so must our assessments. In general, assessment becomes a more integral part of the learning process, growing both broader and deeper to probe student un-

derstanding. It becomes broader in the sense that it encompasses more varied formats of assessment; it is deeper in terms of measuring more complex skills. As students carry out laboratory investigations that challenge them to increase their conceptual understanding, the distinction between assessment and instruction blurs into a seamless whole, and there is near perfect alignment with standards (outcomes and expectations), programs (instruction), and assessments. As we assess scientific thinking, science inquiry, and problem-solving skills, then we must change our instruction to provide students with opportunities to learn and practice these skills.

Figure 1.3 (page 5) depicts a congruence triangle where standards, instruction, and assessment interact in the planning and implementation of successful science programs. If any of the three dimensions does not clearly link or interface with the other dimensions, then we compromise the fairness, credibility, validity, and utility of the assessment. Figure 1.4 (page 5) provides a checklist that teachers and school administrators can apply to evaluate their assessment programs.

Figure 1.3: Congruence Triangle. Reynolds, et al., 1996.

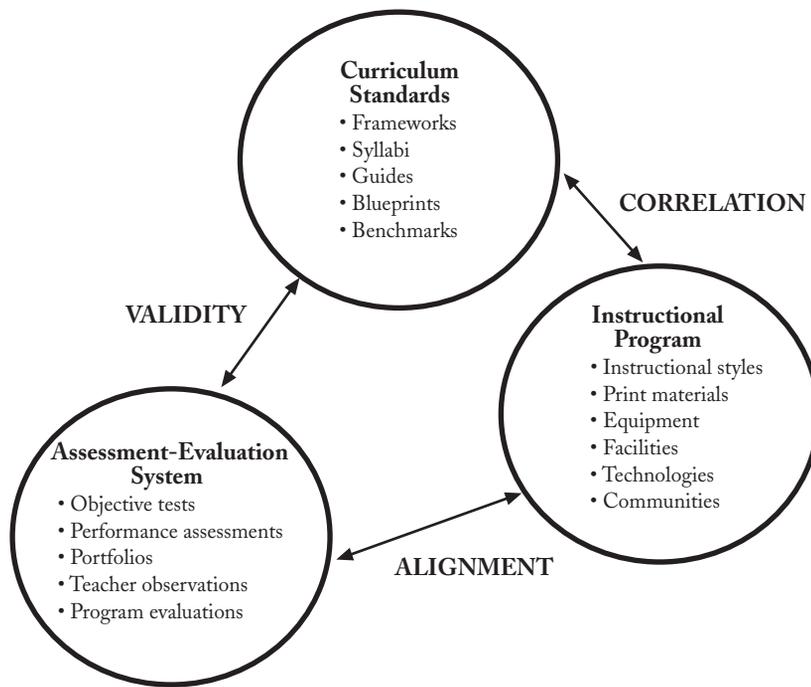


Figure 1.4: Assessment Checklist.

Question	Yes	No
Do the school, district, or state curriculum guides and assessment frameworks incorporate the <i>National Science Education Standards</i> ?		
Are the assessment standards relevant to local perspectives and issues?		
Are the assessment standards developmentally appropriate for the age of students?		
Are the assessment standards challenging to the academic capabilities of students?		
Are the instructional activities of teachers aligned with the assessment standards in use by the school or district?		
Can students distinguish between instruction and assessment?		
Are adequate materials available for student use in the laboratory?		
Are students informed of the criteria for success?		
Are students involved in the development of criteria for success?		
Are the science process skills and content outcomes being measured consistent with the standards in use?		
Do the assessment instruments reflect a variety of formats? Is the assessment system multifaceted?		

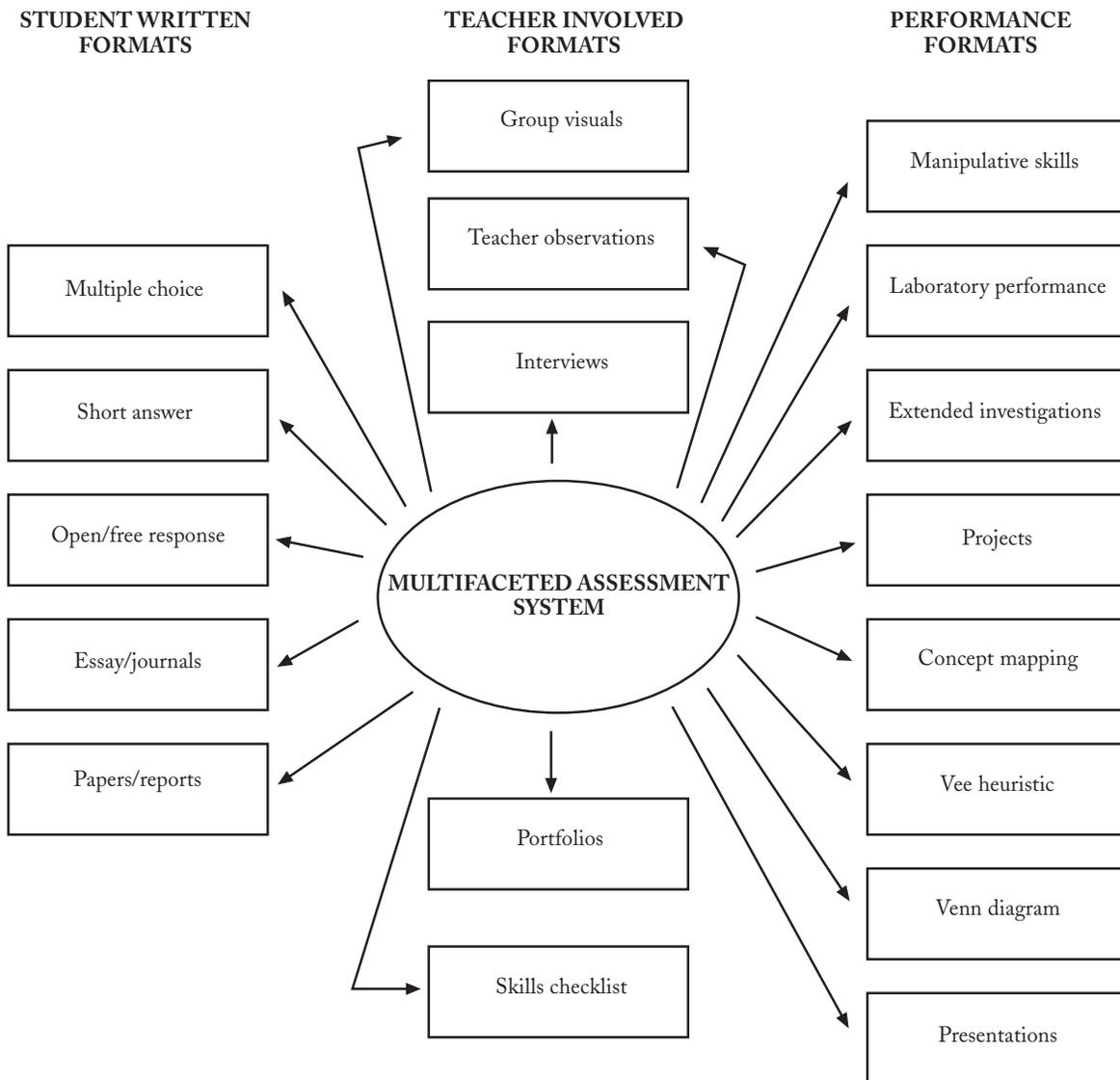
The Multifaceted Assessment System

Educators have traditionally made wide use of paper-and-pencil examinations, which have typically included multiple-choice, true/false, short-answer, and essay questions. Often these assessments are used primarily at the end of a course or instructional unit as a way of measuring overall student understanding of facts and concepts. The large majority of questions in these examinations or assessment formats tend to measure low-level cognitive skills.

With recent reforms, these assessments are being supplemented with a broad range of assessment tools designed to measure higher-level cognitive skills, such as problem-solving, inquiry, communication, and also interpersonal skills. These multifaceted tools can include a variety of assessment formats, as depicted in Figure 1.5.

These varying assessment formats are discussed in greater depth in Chapter 3. They can be used throughout the instructional process to promote student learning.

Figure 1.5: Multifaceted Assessment System. Adapted from Reynolds et al., 1996.



Most of these methods share a common benefit. As you measure student progress during implementation of your science program, you can use the data to adjust instruction and provide assistance to individual students as necessary. The data you collect can also help you adjust overall instructional strategies for use in future science classes.

Teachers can select their most appropriate teaching strategies that help students learn new concepts within the confines of their classroom environment. You can also use the most appropriate assessment formats and techniques to determine whether students have mastered new skills and understandings. Just as no one teaching strategy will cover every learning situation, no single assessment format can measure every aspect of student learning.

The assessment formats depicted in Figure 1.5, for example, are contained within neat little cells. While these formats do provide important data about student learning, in reality a given test might fit into more than one category, or even provide information that supports data gathered by several assessment methods.

This book focuses on performance assessments, and how these assessments connect and interface with the *National Science Education Standards*. Its focus is on performance-based assessments that use the science classroom and laboratory as major contexts for inquiry. Performance-based assessment is by definition “authentic” in nature, because it allows students to demonstrate their science inquiry, reasoning, and understanding skills when challenged with relevant, “real-world”

Figure 1.6: Important Aspects of Laboratory Performance-Based Assessment.

The laboratory is an important component of science instruction.

- There are certain features that are common to all models of laboratory performance-based assessment. There is a Planning and Designing phase or step, a Performing or Doing phase, an Analysis and Interpretation of Data phase, and a Conclusions and Making Projections for Future Study phase. The phases are placed in sequence for discussion purposes. In reality, the phases or steps are interrelated, and students can revisit or retrace their thinking at any time to modify their work or investigation.
 - The laboratory provides an appropriate context for students to engage in problem-based learning, where they practice and use science process and problem-solving skills.
 - Laboratory investigations and tasks by their nature allow students to produce a product and generate, rather than select, responses to questions.
 - If appropriately designed, laboratory investigations allow students to generate multiple solutions to novel problems.
 - As students produce a product and generate multiple responses to questions, laboratory investigations fit the criteria as being performance based.
 - As laboratory performance-based assessment becomes an integral part of science learning, then instruction and the nature of what goes on in science classrooms come closer to the vision of assessment laid out in the *National Standards*. Instruction moves from “a transmission of information” approach to a hands-on, problem-based approach that allows students to integrate new knowledge and skills into their existing cognitive structure.
 - The laboratory or practical science is a “holistic activity” (Woolnough 1991) where students do a task rather than write about something. This in essence is a performance-based activity for a limited or extended period of time. This approach is in agreement with the *National Science Education Standards* for assessment.
 - Laboratory investigations, while an exemplar of performance-based assessment, are also an excellent approach to problem-based learning. Problem-based learning is where students inquire, debate, and engage in discussion of open-ended problems that have multiple solutions. The entire investigation can focus on a single problem.
-

problems. The science laboratory, traditionally under-used as a context for assessment, is an ideal setting for teachers to implement many of the reforms suggested by the *National Science Education Standards*, state assessment frameworks, and other standards documents, such as the New Standards Project (1997a, 1997b). Figure 1.6 (page 7) provides an outline of important aspects of performance-based assessment for the science laboratory.

This conceptualization of science inquiry and its interface with laboratory performance-based assessment is consistent with the assessment standards provided in the *National Science Education Standards*, and forms the basic framework for designing performance assessments.

Many traditional assessments have been large-group oriented—that is, a single teacher administering tests to a class. The new assessment formats supplement these formats by focusing on individuals and small groups. Portfolios, interviews, journals, and other assessment formats reinforce individualized instruction, and also accommodate different learning styles, exceptional students, and students with Limited English Proficiency skills.

Presentations, group and peer evaluations, and projects tap into students' creativity and planning and speaking skills by providing them with the opportunity to do the same things adults do every day. Life is not a series of true/false or multiple-choice tests. In most "real-world" decision-making and problem-solving situations, adults gather appropriate information, interpret that information using their own experiences and knowledge, and reach appropriate conclusions. In many cases, their decisions have important consequences. In the process, adults discard irrelevant information, search for additional data, and anticipate the consequences of their actions.

They also communicate their decisions, along with their rationale, to others.

A significant component of our current teaching and assessment is based on words—transmitting information to students verbally and through print, and then requiring students to repeat or replicate that information verbally and through writing. But many students learn best by receiving information through visual tools such as charts, data tables, graphs, and sketches. For such students, these kinds of visual stimuli can produce more effective learning. Several of these student performance-based assessment formats—including concept maps, Venn diagrams, and the Vee heuristic (see pages 35–42 for examples of all three)—emphasize visual stimuli.

Alternative response formats offer significant assistance to learners with Limited English Proficiency skills and other exceptionalities. As teachers, we must be willing to accept many kinds of evidence given by students to demonstrate their understanding of a concept or principle. As there are many ways to demonstrate understanding, we need to go beyond paper-and-pencil assessment formats and embrace alternative assessment formats that reflect a variety of learning styles, cooperative learning in small groups, and the nurturing of multiple intelligences.

Using Assessment Results—The New Paradigm

Science classroom and laboratory assessments are the foundation of a sophisticated process designed to evaluate and improve the science education system. Everyone—from students, teachers, and parents to government officials—uses assessment data to evaluate how well the education system is performing. It's all part of a growing em-

phasis on making the education system accountable for its progress. According to the *National Science Education Standards*:

Assessment is the primary feedback mechanism in the science education system. For example, assessment data provide students with feedback on how well they are meeting the expectations of their teachers and parents, teachers with feedback on how well their students are learning, districts with feedback on the effectiveness of their teachers and programs, and policymakers with feedback on how

well policies are working. Feedback leads to changes in the science education system by stimulating changes in policy, guiding teacher professional development, and encouraging students to improve their understanding of science.

Figure 1.7 depicts some of the components in the four-part assessment data collection process designated in the *National Standards*, and highlights the complexity of assessment and how different parts all work together to provide a basis for important decisions.

Figure 1.7: Components in the Assessment Data Collection Process. *National Science Education Standards*, NRC, 1996.

The four components can be combined in numerous ways. For example, teachers use student achievement data to plan and modify teaching practices, and business leaders use per capita educational expenditures to locate businesses. The variety of uses, users, methods, and data contributes to the complexity and importance of the assessment process.

Data Use	Data Collection	Collection Methods	Data Users
	To describe and quantify:		
Plan teaching	Student achievement and attitude	Paper-and-pencil testing	Teachers
Guide learning	Teacher preparation and quality	Performance testing	Students
Calculate grades	Program characteristics	Interviews	Parents
Make comparisons	Resource allocation	Portfolios	Public
Credential and license	Policy instruments	Performances	Policymakers
Determine access to special or advanced education		Observing programs, students, and teachers in classroom	Institutions of higher education
Develop education theory		Transcript analysis	Business and industry
Inform policy formulation		Expert reviews of education materials	Government
Monitor effects of policies			
Allocate resources			
Evaluate quality of curricula, programs, and teaching practices			

Conclusion

It is clear that assessment is an important, integral part of science education that promotes learning for all students. Teachers use a variety of assessment instruments of the highest quality for providing feedback to students, parents, administrators, and policymakers. There is no single assessment format that works best for everyone; you must refine your assessments through trial and error to develop a system that works best for your particular situation. Different assessment formats provide different kinds of information used for different purposes. Classroom and laboratory assessments focus on improving student learning by providing feedback to students, while international and national assessments provide data for system accountability.

The next three chapters of this book focus on developing performance assessment tasks, alternative forms of assessment, and the analysis and use of assessment data. These chapters will give you a practical primer on how to improve the assessment process in your classroom or school.

Works Cited

- National Research Council (NRC). 1996. *National Science Education Standards*. Washington, DC: National Academy Press.
- News Standards Project. 1997a. *Performance Standards. Volume 2: Middle School*. Washington, DC: National Center for Education and the Economy (Tel. 202-783-3668).
- . 1997b. *Middle School Science Portfolio*. Washington, DC: National Center for Education and the Economy (Tel. 202-783-3668).
- Perkins, D., and Salomon, G. 1989. Are Cognitive Skills Context-Bound? *Educational Researcher* 19:16–25.

Reynolds, D., Doran, R., Allers, R., and Agruso, S. 1996. *Alternative Assessment in Science: A Teacher's Guide*. Buffalo: University of Buffalo.

Rubba, P., Miller, E., Schmalz, R., Rosenfeld, L., and Shyamal, K. 1991. Science Education in the United States: Editors Reflections. In *Science Education in the United States: Issues, Crises and Priorities*. Easton, PA: Pennsylvania Academy of Science.

Suggested Readings

- Carr, M., Barker, M., Bell, B., Biddulph, F., Jones, A., Kirkwood, V., Pearson, J., and Symington, D. 1994. The Constructivist Paradigm and Some Implications for Science Content and Pedagogy. In *The Content of Science—A Constructivist Approach to Its Teaching and Learning*, Fensham, P., Gunstone, R., and White, R., eds. Bristol, PA: Falmer Press.
- Duit, R., and Treagust, D. 1995. Students' Conceptions and Constructivist Teaching Approaches. In *Improving Science Education*, Fraser, B., and Walberg, H. eds. Chicago: National Society for the Study of Education.
- National Center on Education and the Economy, University of Pittsburgh. 1997. *Performance Standards, Volumes I, II, and III*. Washington, DC: National Center on Education and the Economy.
- National Research Council. 2000. *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*. Washington, DC: National Academy Press.
- New York State Education Department, University of the State of New York. 1996. *Learning Standards for Mathematics, Science, and Technology*. Albany: New York State Education Department.
- Woolnough, B. 1991. Practical Science as a Holistic Activity. In *Practical Science*, Woolnough, B. ed. Bristol, PA: Open University Press.

Yager, R. 1995. Constructivism and the Learning of Science. In *Learning Science in the Schools: Research Reforming Practice*, Glynn, S., and Duit, R., eds. Mahwah, NJ: Lawrence Erlbaum Associates.