College Science Teachers Guide to Assessment

Editors
Thomas R. Lord
Donald P. French
Linda W. Crow

NSTA press
National Science Teachers Association
Arlington, Virginia

NSTA press®
National Science Teachers Association

Claire Reinburg, Director
Jennifer Horak, Managing Editor
Judy Cusick, Senior Editor
Andrew Cocke, Associate Editor
Betty Smith, Associate Editor

Art and Design
Will Thomas, Jr., Director
Joseph Butera, Graphic Designer, cover and interior design

Printing and Production
Catherine Lorrain, Director
Nguyet Tran, Assistant Production Manager
Jack Parker, Electronic Prepress Technician

National Science Teachers Association
Francis Q. Eberle, PhD, Executive Director
David Beacom, Publisher

12 11 10 09 4 3 2 1

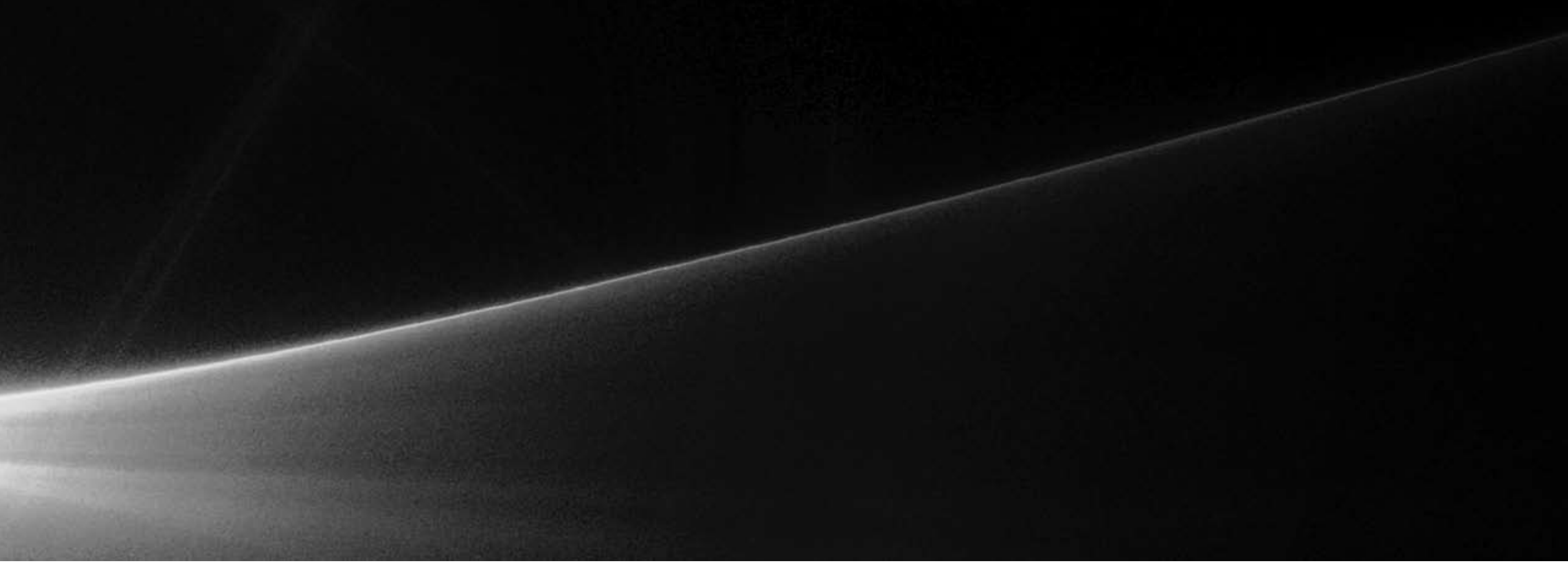
Library of Congress Cataloging-in-Publication Data

College science teachers guide to assessment / Thomas Lord, Donald French, and Linda Crow, editors.
p. cm.
Includes index.
ISBN 978-1-933531-11-3
1. Science—Study and teaching (Higher)—Methodology. 2. Science—Study and teaching (Higher)—Evaluation. 3. Academic achievement—Evaluation. 4. Educational tests and measurements. I. Lord, Thomas R. II. French, Donald P., 1952- III. Crow, Linda W.
Q181.C535 2009
507.1'1—dc22
2008041750

NSTA is committed to publishing material that promotes 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 authors do not warrant or represent that the procedures and practices in this book meet any safety code or standard of federal, state, or local regulations. NSTA and the authors 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.

Table of Contents

Table of Contents



Section 3. How-To Section
Successful Classroom-Tested Practices and Instructions and Rubrics for Their Implementation

Table of Contents

Section 4. General Practices to Improve Assessment

Contributors

Brown, Mary H.
Lansing Community College
MC 5400–Science
411 N. Grand Ave.
Lansing, MI 48901
Brownm@lcc.edu

Cheesman, Kerry L.
Capital University
Columbus, OH
kcheesma@capital.edu

Coleman, Anne
Department of Life and Physical Sciences
Cabrini College
610 King of Prussia Rd.
Radnor, PA 19087
annecoleman@cabrini.edu

Cohen, Robert A.
Department of Physics
East Stroudsburg University
East Stroudsburg, Pennsylvania
rcohen@po-box.esu.edu

Cudaback, Cynthia
Marine Earth and Atmospheric Science
Campus Box 8208
North Carolina State University
cynthia_cudaback@ncsu.edu

Dean, Karol
Mount St. Mary's College
12001 Chalon Rd.
Los Angeles, CA 90049
kdean@msmc.la.edu

Eason, Grace
University of Maine at Farmington
173 High St., Preble Hall
Farmington, ME 04938
geason@maine.edu

Gallo, Mark A.
Niagara University
DePaul Hall,
Niagara University, NY 14109
mgallo@niagara.edu

Gomez-Zwiep, Susan
California State University, Long Beach
Department of Science Education
Long Beach, CA 90840
sgomezwp@csulb.edu

Haviland, Don
California State University, Long Beach
1250 Bellflower Blvd.
Long Beach, CA 90808-2201
dhavilan@csulb.edu

Kandlbinder, Peter
Institute for Interactive Media and Learning,
University of Technology, Sydney
Peter.Kandlbinder@uts.edu.au

Lord, Thomas R.
Department of Biology
Indiana University of Pennsylvania
Indiana, PA 15705
trlord@iup.edu

Melvin, Thomas J.
Biology Department
Indiana University of Pennsylvania
Indiana, Pennsylvania 15705
t.j.melvin@iup.edu

Metz, Anneke M.
Department of Cell Biology and Neuroscience
Montana State University
Bozeman, MT 59717
anneke@montana.edu

Contributors

Rybarczyk, Brian J.
The Graduate School
University of North Carolina/Chapel Hill
CB#4010
Chapel Hill, NC 27599-4010
brybar@unc.edu

Siebert, Eleanor D.
Mount St. Mary's College
12001 Chalon Rd.
Los Angeles, CA 90049
esiebert@msmc.la.edu

Sorensen, Kathryn H.
Biology Department
American River College
4700 College Oak Dr.
Sacramento, CA 95841
SorensKH@arc.losrios.edu

Straits, William J.
California State University, Long Beach
Department of Science Education
Long Beach, CA 90840
wstraits@csulb.edu

Swails, Nancy J.
Capital University
Columbus, OH 43209
swails@capital.edu

Tichenor, Linda
Department of Biology
University of Arkansas
Fort Smith, Arkansas 72913
lticheno@uafortsmith.edu

Travis, Holly
Department of Biology
Indiana University of Pennsylvania
Indiana, Pennsylvania 15705
holly.travis@iup.edu

von Bergmann, HsingChi
EDT 726, Faculty of Education
University of Calgary
Calgary, AB, Canada T2N 1N4
hsingchi@ucalgary.ca

Waldvogel, Jerry A.
Department of Biological Sciences
Clemson University
Clemson, SC 29634-0314
waldvoj@clemson.edu

Walton, Kristen
Department of Biology
Missouri Western State University
4525 Downs Dr.
St. Joseph, MO 64507
kwalton1@missouriwestern.edu

Wilke, R. Russell
Angelo State University
Department of Biology
ASU Station - 10890
San Angelo, TX 76909
russell.wilke@angelo.edu

Wood, Bonnie S.
University of Maine at Presque Isle
181 Main St.
Presque Isle, ME 04769
bonnie.s.wood@umpi.edu

Yerger, Ellen
Department of Biology
Indiana University of Pennsylvania
114 Weyandt Hall
Indiana, Pennsylvania 15701
ellen.yerger@iup.ed

Preface: Note From the Editors

Welcome to the sixth in the series of SCST monographs created by the members of the Executive Board of the Society for College Science Teachers in cooperation with the National Science Teachers Association. This document covers an extremely important and often controversial topic, that of evaluating the value of students and professorial works. The jointly sponsored monograph has been three years in the making with the initial agreement with the National Science Teachers Association taking place in the fall of 2006.

Each submission in this monograph was reviewed by at least two members on the Editorial Board with the published authors responding to reviewers critiques and providing the final proofing of their own entry. Articles were selected on the quality of the writing and their contribution to the value and importance of assessment in a college science setting.

The monograph examines assessment issues from several different viewpoints and is broken into several chapters. The first section deals with general assessment topics such as validation of survey instruments and creating a culture for faculty-owned assessment. The second section concerns traditional and alternative forms of assessment in both science and the science education classroom. The third section presents a series of how-to assessment practices that have been successfully utilized in the field. Finally, the fourth section provides a series of tips to enhance assessment in the college science classroom.

The editors would like to thank all the contributors to the monograph. The quality of the initiative is indicative of the time and energy they put into this work.

Thomas R. Lord
Indiana University of Pennsylvina
trlord@iup.edu

Donald P. French
Oklahoma State University
dfrench@okstate.edu

Linda W. Crow
Montgomery College
linda.w.crow@nhmccd.edu

Acknowledgments

The editors wish to thank Holly Travis for her tireless effort and dedication to the construction of this document. Thanks also to Ellen Yerger and Tom Melvin for their help in making this monograph a success.

Chapter 5

Writing/Using Multiple-Choice Questions to Assess Higher-Order Thinking

Kerry L. Cheesman
Biological Sciences
Capital University
Columbus, Ohio

Introduction

Multiple-choice exams are widely used in college science classrooms (as well as for laboratory quizzes and exams). Multiple-choice questions have many advantages—perhaps the most important is that they can be graded quickly and easily, and they can be graded by either human or machine. The "clicker" systems often used in large lecture rooms are well adapted for answering multiple-choice questions, and they can be used for "instant quizzes" with immediate feedback to students.

Instructor time is valuable, and in large classrooms the use of essay exams (the primary alternative) can quickly become overwhelming, causing students to wait for feedback for prolonged periods of time. Feedback on progress needs to be as rapid as possible, and essay questions do not lend themselves to that. Essay grading can also tend to be biased by any number of factors (time of day, personal biases, differences between graders, lack of openness to new interpretations, and so on).

Finally, most graduate entrance exams (including the GRE, MCAT, and DAT) are based on multiple-choice questions. Many later exams, such as the medical board exams, are also multiple-choice. Therefore, it is important to make sure that students are prepared for higher-order multiple-choice exams and the reasoning that is required to answer the questions in a proficient manner. Undergraduate science instructors can help students be well prepared by using higher-order multiple-choice questions for assessment of course material starting in the freshman year.

Assessment must match one's teaching style—inquiry teaching must be followed by assessment techniques that match the inquiry method of teaching. If one follows the learning cycle (5E or other similar models), assessment is encountered throughout the teaching and learning continuum, and that assessment must be related to the phase of the cycle (exploration, extension, etc.). Certainly higher-order questions capture the essence of exploration and extension much better than lower-order questions do.

The use of higher-order questions does not mean an end to using lower-order questions. Rather, we are referring to a shift from the t 80–90% lower-order questions typically found in college science exams toward a balance between lower- and higher-order questions. The goal of undergraduate science instruction should be critical thinking rather that memorization. Many students come to the university with the assumption that science is just a lot of memorization, and college instructors often need to work hard to destroy that myth. However, that myth is often kept alive by the choice of questions used on the exams. If they favor knowledge-style questions, then students will continue to believe that science is mostly about memorization rather than about inquiry and analysis.

Understanding Bloom's Taxonomy

Most college instructors are familiar, on some level, with Bloom's taxonomy of learning (Bloom et al. 1956). Much has been written about the use of Bloom's taxonomy in the construction of exam questions, but few instructors take to heart the need to use all of the levels instead of just the first two in constructing examination questions. Here is a quick review of Bloom's taxonomy as it relates to the teaching of college science.

1. Knowledge: the ability to remember/recall previously learned material.

Examples of behavioral verbs: list, name, identify, define, show

Sample learning objectives in science: know common terms, know specific facts, know basic procedures and methods

2. Comprehension (understanding): the ability to grasp the meaning of material, and to explain or restate ideas.

Examples of behavioral verbs: chart, compare, contrast, interpret, demonstrate

Sample learning objectives in science: understand facts and principles, interpret charts and graphs, demonstrate laboratory methods and procedures

3. Application: the ability to use learned material in new situations.

Examples of behavioral verbs: construct, manipulate, calculate, illustrate, solve

Sample learning objectives in science: apply concepts and principles to new situations, apply theories to practical situations, construct graphs and charts

4. Analysis: the ability to separate material into component parts and show relationships between the parts.

Examples of behavioral verbs: classify, categorize, organize, deduce, distinguish

Sample learning objectives in science: distinguish between facts and inferences, evaluate the relevancy of data, recognize unstated assumptions

5. Synthesis: the ability to put together separate ideas to form a new whole or establish new relationships.

Examples of behavioral verbs: hypothesize, create, design, construct, plan

Sample learning objectives in science: propose a plan for an experiment, formulate a new scheme for classifying, integrate multiple areas of learning into a plan to solve a problem

6. Evaluation: the ability to judge the worth or value of material against stated criteria.

Examples of behavioral verbs: evaluate, recommend, criticize, defend, justify

Sample learning objective in science: judge the way that conclusions are supported by the data

It is a common misconception that as one climbs the scale of Bloom's taxonomy, the difficulty of the questions increases. The increase in cognitive demand associated with higher-order questions refers to the *complexity* of the questions, not the difficulty. Higher-order questions require a different set of cognitive demands, but they are not necessarily more difficult.

Writing Multiple Choice Questions

Higher-order multiple-choice questions can be as easy or as difficult to construct as lower-order questions. Good-quality questions are essential to being able to truly assess a student's knowledge and understanding of the subject matter in any area of science.

Before attempting to construct individual questions, think about the purpose of the questions. In general, the purpose should be to assess what students know and don't know, and how students are able to construct knowledge based on prior learning. Therefore, avoid "trick" questions that may confuse students who understand the material. Avoid using prepared test banks written by the author of the textbook or other contracted writers. Honest assessment *must* match the teaching style employed, not the style of the textbook or the style of your colleagues. Note: You cannot ask higher-order questions if your teaching style mandates only recall.

Writing good multiple-choice questions takes time—a well-constructed test can't be written in a single day. Questions need to be written, reviewed for clarity, and often revised. Questions need to be constructed in such a way that they neither reward test-wise students nor penalize those whose test-taking skills are less developed. The purpose is to assess student learning, and therefore each question needs to be clearly designed to achieve that goal. Remember that higher-order questions take longer to answer than recall questions, so plan accordingly in the construction of the test.

To construct a higher-order multiple-choice question, start by constructing the stem. The stem should pose a problem or state a question. Familiar forms include case study, premise and consequence, problem and solution, incomplete scenario, and analogy. The stem may involve pictures and diagrams or just words.

Write the stem as clearly and simply as possible. A student should be able to understand the problem without having to read it several times. Always try to state the problem in a positive form, as students often misread negatively phrased questions. Avoid extraneous language that is irrelevant to the question. While some authors believe this helps separate those who truly understand from those who don't, too often it confuses even the well-prepared students, leading to unreliability of the question.

Never use double negatives. Avoid "which of these is the *best* choice" unless that format is integral to the learning objectives. Be sure to include in the stem any words that are redundant to all of the answers, and use "a(n)" or "a/an" to avoid eliminating any of the answers as mismatches.

Once the stem is constructed, proceed with writing the responses. Write the correct answer first. This allows you to be sure it is well constructed and accurate, and allows you to match the remaining answers to it. Avoid verbal cues, and certainly avoid lifting phrases directly from the text or class notes. Be sure that the incorrect responses match the correct one in length, complexity, phrasing, and style. For instance, in the following example, the mismatch of the answers makes it easy to guess the correct response even if one has little knowledge of the subject material.

The term "side effect of a drug":

a. refers to any action of a drug in the body

other than the one the doctor wanted the drug to have
b. is the main effect of a drug
c. additionally benefits the individual

Distracters (incorrect answers) must be incorrect yet plausible. If a recognizable key word appears in the correct answer, it should appear in some of the distracters as well. Be sure to check—will the answers help to distinguish *why* a student got it wrong? This is an important part of assessment that is often overlooked by instructors, but is a critical part of helping students to learn.

Avoid microscopic distinctions between answers, unless this is a significant objective of the course. Be sure to stagger the correct responses in their order (use all answer positions as equally as possible). Limit the number of options—most authors agree that 4–5 answers is plenty, and there is no assessment advantage in using more than five. Use *all, always, never, none,* etc., rarely. These are answers that students have been programmed to shy away from and may distort the question as a valid assessment tool. Likewise, use *all of the above* and *none of the above* sparingly.

When all exam questions have been constructed, check each one to see where it falls in Bloom's hierarchy. Construct a simple table such as that shown in Table 1 to see the distribution of questions. If the questions are disproportionately distributed, then rewrite enough questions to balance the exam between lower-order and higher-order questions.

Examples of Multiple-Choice Questions at Each Level

The following examples illustrate the construction of multiple-choice questions that fit the higher levels of Bloom's taxonomy. For most an explanation is included describing why it fits where it does, and what a student needs to know to be able to answer the question correctly.

TABLE 1. Sample table showing the distribution of questions

Topics/ Objectives	Recall (knowledge, comprehension)	Application (application, analysis)	Evaluation (synthesis, evaluation)	%
a.				
b.				
c.				
d.				
e.				
f.				
Total				

Application Questions

1. Susie and Bill are both healthy and have healthy parents. Each of them has a sister with *autosomal recessive cystic fibrosis*. If Susie and Phil have a child, what is the probability that it will be born with cystic fibrosis?
   a. 0
   b. 2/3
   c. 1/2
   d. 1/4
   e. 1/9
   (To answer this question correctly, one must understand the terms *autosomal* and *recessive*, and also understand the concepts of probability as applied to human genetics. In this question the student must apply those concepts to a family situation not studied before. The incorrect responses are constructed to find misconceptions/misunderstandings about genetic probabilities.)

2. A total of 100 students at Capital University were tested for blood type. The results showed 36 were type O, 28 were type A, 28 were type B, and 8 were type AB. The frequency of the A allele is therefore:
   a. 0.10
   b. 0.14
   c. 0.28
   d. 0.56
   e. 0.64
   (To answer correctly, a student must know the formula for allele frequency and be able to calculate it from the given data. The answers were chosen to help find misunderstandings about allele frequency.)

3. Evolutionary forces have produced an unusual plant, the Indian Pipe, that has no chlorophyll. Therefore, the plant must:
   a. make its own food
   b. absorb food made by other organisms
   c. photosynthesize without chlorophyll
   d. respire without taking in food
   e. use chlorophyll from other plants
   (To answer this question the role of chlorophyll in energy transformation must be understood. The student must apply the concepts of energy transformation/lack of chlorophyll to a logical new endpoint. All of the answers are plausible and help to distinguish where an understanding of energy transformation is incomplete.)

4. Which of the following compounds should have the highest boiling point?
   a. $CH_3CH_2CH_2CH_3$
   b. CH_3NH_2
   c. CH_3OH
   d. CH_2F_2
   (To answer correctly, a student must understand the concept of boiling point and the role of various constituent chemical groups in raising or lowering the boiling point.)

Analysis Questions

1. When a solid ball is added to a graduated cylinder with water in it, the water level rises from 20 ml to 50 ml. What is the volume of the ball?
   a. 20 ml
   b. 30 ml
   c. 50 ml
   d. 70 ml
   (In this example, a student must understand the concept of volume and not get distracted by the spherical nature of the added object. The answers are designed to give the instructor a sense of the misunderstanding of volume.)

2 From the graph shown here, determine when maximal carrying capacity has been reached.
   a. point A on the graph
   b. point B on the graph
   c. point C on the graph
   d. point D on the graph
   (To answer correctly a student must be able to see the relationships among the various organisms, interpret those relationships, and evaluate the ecosystem relative to the organisms shown on the graph.)

3. During an otherwise normal pregnancy, a woman begins to experience light-headedness and a decline in energy levels near the end of the first trimester. Which of the following is the most likely cause of her symptoms?
   a. lack of B vitamins due to poor diet
   b. decline of blood pH due to overuse of muscles
   c. decrease in blood pressure due to expanding fetal circulation
   d. decline in estrogen levels due to ovarian shutdown

   (All of these answers involve factors that could cause tiredness in a woman. To determine the most likely cause in this scenario, a student needs to understand the basic mechanics of pregnancy and the biochemical changes that occur during it. The answer given shows a student's ability to carefully analyze the situation and determine causality.)

4. The seeds of various plants vary in size from a fraction of a millimeter to several centimeters. The most critical factor controlling the size seed a plant produces is
   a. size of the maternal flower
   b. projected size of the animal pollinator
   c. quantity of the abiotic pollinator
   d. length of predicted dormancy
   e. method of distribution of seed

Synthesis Questions

1. Domoic acid, produced by diatoms, has been found to bind to hippocampal glutamate receptors. If a person were to accidentally consume a lot of shellfish contaminated with this organism, what effect might be expected?
   a. blindness
   b. deafness
   c. amnesia
   d. aphasia
   e. rigidity

   (To answer this question, one must understand the role of the hippocampus and the role of glutamate in this area of the brain. Here a student's knowledge of the brain is used to establish a new relationship beyond those studied in class. All of the answers are logical outcomes of brain dysfunction and help the instructor to pinpoint the misunderstandings that students have.)

2. A neighbor found some mammal bones on the ground—a keeled sternum and an ulna, where the olecranon occupies 30% of the length of the bone. These bones most likely came from what type of mammal?
   a. flyer
   b. climber
   c. runner
   d. digger
   e. swimmer

   (Here a student must understand the various bones and what their functions are. The student must then formulate a relationship between the type and formation of the bones and the activity that it would promote in a mammal.)

3. What would be the most logical result of mixing *X* and *Y*, both solubilized in distilled H_2O at room temperature?
   a. precipitation of a solid
   b. a change in color of the liquid
   c. a rapid rise in temperature
   d. a rapid decrease in temperature

   (To answer correctly, a student must put together knowledge of *X* and *Y* as compounds with knowledge of their dissociations and the reactions of the individual components. All of the answers reflect outcomes that the student has previously experienced when two compounds are mixed.)

Evaluation Questions

1. Your fitness regimen involves jogging on the school track 2–3 miles per day with a friend. On a particular day, about 15 minutes into your jog, your friend suddenly pulls up and falls down, grasping her right calf in pain.

What should you do at that moment?

   a. apply ice to the calf
   b. apply heat to the calf
   c. tell her to get up and walk slowly
   d. get emergency help stat

(Here the student must be able to appraise the situation and evaluate the next course of action. The student's knowledge of both muscle function and injury are brought to bear in deciding which treatment to use.)

2. According to the American Heart Association, obesity plays a major role in early heart failure. Which of the below answers best describes how being severely overweight can cause the heart to falter?
   a. Obesity creates the need of greater blood volume in the circulatory system.
   b. Obesity creates the need for a quicker heart rate each minute.
   c. Obesity creates the need for a larger volume of blood exiting the heart per beat.
   d. Obesity causes the heart to grow larger than is anatomically recommended.
   e. Obesity causes the heart to develop a greater cardiac vessel network.

3. Abiotic factors impact heavily on photosynthetic richness in a green plant. The basic factors that influence sugar making in plants are (1) quantity of water, (2) quantity of sunlight, (3) quantity of CO^2, (4) environmental temperature, (5) movements of the air, and (6) richness of growth substrate. By moderately increasing in quantity, which of the factors would positively influence photosynthesis and which factors would negatively influence photosynthesis?
   a. + (1), (2), (3) and – (4), (5), (6)
   b. + (2), (3), (5), (6) and – (1), (4)
   c. + (1), (2), (3), (4), (5) and – (6)
   d. + (3), (4) and – (1), (2), (5), (6)
   e. + (1), (2), (3), (4) and – (5), (6)

Conclusion

Multiple-choice questions may be used effectively to assess student learning as long as they are constructed properly and include assessment of higher-order thinking. Taking the time to construct good stems and good answers, not only on the exams but also in the daily questions posed during the lesson, is well worth the effort. Practicing coming up with conclusions for critical-thinking questions is as important in creating high-level thinking in students as designing good-quality multiple-choice exams. Asking only non-challenging thinking questions during class is a waste of both the instructor's and students' time and does little to help assess student learning.

Reference

Bloom, B., M. Englehart, E. Furst, W. Hill, and D. Krathwohl. 1956. *Taxonomy of educational objectives: The classification of educational goals*. New York: McKay Publishers.

Chapter 24

Eleven Assessment Lessons Learned at the Gate

Mary H. Brown
Lansing Community College
Lansing, Michigan

As a "gateway" instructor for more than 30 years, I've learned a few things about assessing the "typical" community college student. "Gateway" at my institution is the polite euphemism for suggesting you'll always be teaching the nonscience majors with the slim hope that some may eventually learn to tolerate the subject. It's been nearly 20 years since I've evaluated one true science discipline major among the hundreds of students in my classes each academic year. As STEM students are nonexistent in my classes, the best I can hope for is the integrated science major in education.

"Typical community college student" is an oxymoron. There isn't one. Some of my college freshmen are older than I am, returning to college for an opportunity at a retirement career. A few freshmen are dual-enrolled high school students, 16 or 17 years old. Some of my students are parents of young children, and some are grandparents. Today's community college students come from all walks of life and include working adults and recent immigrants or refugees from foreign lands. Included in this mixture is the university's student of choice, whose parents chose the inexpensive route for their first two years of college.

So, how does a "gateway" instructor assess the learning of the "typical" community college student? Included in this chapter are 11 assessment lessons I've learned along the journey.

1. Assessments need to be frequent and scaffolded for legitimate success. Working adults don't want only one or two assessments of their progress during the semester. They want to know exactly how well they are performing in the class at each assignment. They want feedback, personal and directed. These are practical folks! Not surprisingly, they want good grades, but they also want real success. They don't mind being challenged, especially after they've been successful. Successful gateway instructors know that a 16-week semester might have seven or eight large exams. The first exam should be the least challenging. Each subsequent exam should be more challenging. Students want their efforts to show. They want to believe they're progressing, that hard work pays off.

   The experts might suggest that assessments be formative, giving frequent feedback toward the mastery of content. Classroom assessment techniques (CATS) are usually formative assessments, and might include the quick "think pair share" or the "one-minute paper" (Angelo and Cross 1993). Students want definite feedback on all formative assessments. They want to know that you've sincerely read each and every one. Formal grades aren't required, just your attention and constructive remarks in some format.

   Summative assessments are aligned with the evaluation of content mastery or the completion of instruction. Many community college students need state or national benchmarks (standards) for their instruction in their trade or vocational courses. My students see the state and national benchmarks for science education (Roseman and Koppal 2006), and they know that like other professionals, they need to meet those standards. Unlike their certification exams for careers, my class is only the beginning of their journey toward scientific literacy. Benchmarks and standards are a goal for attainment with the expected outcome of lifelong scientific literacy.

2. Assessments and exams are not always the same thing. Assessments come in many forms. In the assessment report I file each year to the divisional office to show that my nonscience majors course is worthy of the title of science "CORE"(which means it meets the criteria of inquiry, shows the processes and limitations of scientific thought, and analyzes data), no fewer than 12 different assessment techniques are listed.

   Assessments include the exams, concept maps (two varieties), Vee diagrams, laboratory reports, and capstone projects. All of these give the instructor information about the students' learning and their mastery of content. Alternative assessments that are real, targeted to the content can be more revealing than an exam. Anxiety plays a role in exam taking, but a student has control over a project. Presentations or projects that allow for research, sharing of ideas, and collaboration are valid assessments. These include contextual, problem-, case-, or performance-based assessments.

3. Embedded assessment across multiple sections has advantages and disadvantages. Community colleges are notorious for having large numbers of adjunct professors. Mine is not an exception. Subsequently, as the full-time professor responsible for reporting on multiple course offerings—even when I am not the "instructor of record"—my job becomes very challenging. Embedded questions on each exam allow for a logistically simple method for tracking all sections of a single course. Instructors simply provide me detail of the embedded questions after each exam. That's the advantage. The disadvantage is that I have no idea why the students miss the embedded questions on particular topics. The variables are too numerous. I have a vague idea within my own classes, as I can monitor absences or recall the day in class when the topic was discussed. There is no information from the classes I didn't teach. The number is a cold sta-

tistic without any qualitative information. After a period of time, the embedded questions on exams must change in their wording. Since the statistical report depends on data from the previous year, altering the question requires a whole set of rationale, without qualitative information. That's another disadvantage.

4. Listening provides more assessment value than talking. As my students engage in their laboratory activities, I listen. I listen to their interactions, collaborations, and arguments. I walk to each lab table, and I listen. I can learn a tremendous amount about their learning and their assimilation of the content by listening. During the course of the semester, they become very accustomed to me walking to each table without saying anything as they work. I learn a lot about their thinking processes by listening. Each unit of instruction also begins with me listening. Each collaborative group is asked to list prior knowledge about topics in the unit. Together we summarize. We use the prior knowledge expressed in discussions to increase the depth of knowledge on the topics. Pre-assessment starts with listening.

   Each unit of the courses I teach starts with a series of connection questions. What do you already know about the topic? How can learning this information be useful to you? What are you looking forward to learning about this topic? Each unit also ends with reflection. What did you learn about this topic? Did anything you learned in this unit change your mind? How will this information be useful for your future? Postassessment also starts with listening. I've learned that students do not necessarily answer these questions unless they are explicitly asked. If there exists a possibility that you could be called upon and expected to respond directly to a specific question, you prepare a response. Without that potential accountability, it's a rare student who prepares a response or who is introspective without the prompting questions.

5. A wrong answer has tremendous value. A well-thought-out, detailed wrong answer gives you lots of information. It provides you opportunity to correct a misconception or to craft a discrepant event to allow the learner to construct a more scientifically accurate response. Providing the question in advance and allowing two minutes to think before calling on a student to respond yields more information than simply calling on a student. Wait time also works well (Rowe 2003). Giving students a 30-second warning before expecting a response is very powerful. "I'm going to ask <student's name> to respond to the next question" gives that individual a few extra seconds to compose an answer. The responses are more complete, even when they are wrong.

6. Assessment need to be clearly tied to outcomes, objectives, or learning targets. Both instructor and student need to clearly know the purpose of the assessment. What are we evaluating? Communications of expectations are important. Providing the format of the exam gives students an opportunity to prepare appropriately. You study differently for a written essay than a multiple-choice exam. You prepare differently for a presentation than for a discussion.

7. Assessments that are viewed as "products" by students are sources of pride. I have many "product" assessments in my classes. I'm always surprised by college students who have told me that their perfect score concept map was hung on their refrigerator! Or the lab report with the phrase "Well done!" was read over the dinner table. It seems it doesn't really matter how old we are, a well-done product is a source of pride. Community college students know the rewards of hard work. They work tremendously hard in the challenges of everyday life. They can do exceptional work when the assessment is viewed as a product.

8. Detailed constructive feedback on assessments is essential. It takes about six hours to correct a stack of 24 lab reports, if they are well writ-

ten. A set of poorly written lab reports takes about twice as long. Each report needs carefully worded constructive feedback. Rubrics on written assignments (and oral for presentations) are given in advance, and students are expected to follow the same criteria for excellence in writing as they would in a composition class. Even on exams, common mistakes are explained. The feedback on the exam is another opportunity to teach.

9. Owning and expressing your expectations for their success is crucial. Students rise to the challenge of high expectations. When given rationale for a challenge, they accept. They will even accept the frustrations of disequilibrium if they understand the rationale. Explicit reasons for content expectations are essential. Community college students will accept the "because it's on the test" but are likely to ask you why it's on the test. They want a more practical reason for learning the content. Ideally, the reason is tied to a potential career, or an everyday application.
10. Assessments need to be varied, perceived as fair and attainable, and evaluated both objectively and subjectively (Mintzes, Wandersee, and Novak, 1999). The brain loves novelty. It fatigues when offered routine. With 12 different types of assessments throughout the semester, fatigue is more physical than cerebral! Each exam has a variety of question types. Students create or correct concept maps; they evaluate true and false statements, correcting the false. They also write brief answers and traditional multiple paragraph essays. Each exam also has multiple-choice questions and paragraph completions. Students analyze their exam results at the conclusion of each unit and write goals to improve weak performance areas. Students know how exams are evaluated. They know that each section is evaluated independently without my knowledge of the test author. They also know when the sections of the exam are totaled, I often write encouraging remarks on their progress (e.g., "Nice improvement on this multiple-choice section; keep working!") Statistical analysis is given on the entire class performance, and the class discusses improvement strategies for the next exam. Besides the exams, alternative assessments are a near daily occurrence.
11. Ideally, assessments inform teaching, and self-assessments can even inform the learner. Assessment is *not* only about evaluating the learning process. It should change the teaching process. Each assessment should inform the instructor as to needed changes in pedagogy, presentation or missing fundamentals for conceptual understanding. Self-assessments can provide the learner with great potential to change.

The view from the "gate" as I encourage students to consider the science disciplines is generally positive. Together, the students and I investigate, listen to each other, and plan our journey together. Although assessment reports are needed for multiple levels (divisional, program, departmental) there is enough consistency across the requirements that only the perspective changes. Not only have I learned how to assess my "typical" community college students so that I know what they are learning, I've also learned how to teach better science through our shared assessments.

References

Angelo, T. A., and K. P. Cross. 1993. *Classroom assessment technique: A handbook for college teachers.* 2nd ed. San Francisco, CA: Jossey Bass Publishers.

Mintzes, J., J. Wandersee, and J. Novak, eds. 1999. *Assessing science understanding: A human constructivist view.* San Francisco, CA: Academic Press.

Roseman, J. E., and M. Koppal. 2006. Ensuring that college graduates are science literate:

Implications of K-12 benchmarks and standards. In *Handbook of College Science Teaching*, eds. J. Mintzes, and J. Leonard, 325–349. Arlington, VA: NSTA Press.

Rowe, M. 2003. Wait-time and rewards as instructional variables, their influence on language, logic, and fate control: Part one—wait-time. *Journal of Research in Science Teaching.* Vol. S1, 19–32.

Index

*Page numbers in **boldface** type refer to figures or tables.*

D

E

F

T

V

W