many different learning strengths. In this case the instructor might consider an assignment "menu" from which students may choose the assignments that best fit their particular learning abilities (APA Board of Educational Affairs 1995). The lesson here is that science should be relevant and accessible to everyone. Research shows that more students learn more science when they see the relevance of science to their own lives and when their learning strengths are addressed (Woolfolk 1998).

From the Field

Literature-Based Examinations and Grading Them: Well Worth the Effort

Brian P. Coppola, Department of Chemistry
University of Michigan, Ann Arbor

Examinations, probably more than anything else, transmit our learning agenda to students. In our course, organic chemistry is structured so that state-of-the-art information from the primary literature used as case studies can be presented to novice students on examinations. This assures us that we are true to the facts of science and not simply inventing trivial derivatives of classroom examples. On the examinations, we include the citation along with some contextualizing statements. We are thus sending two messages to our students: (1) memorizing previous examples is not enough, and (2) understanding the subject matter of the introductory course allows the student to understand what chemists actually say about what they study. We reinforce the idea of multiple representations for the same phenomenon by asking students to provide words, pictures, graphs, and numerical versions of the same idea. On nearly every exam, students suggest unanticipated but completely reasonable alternative solutions. These are important to note in class.

There are a few aspects of our grading practice that are worthy to note.

- *Make improvement count.* Because students develop their new skills at different rates, and because the course is truly cumulative, we have devised ways to make improvement count. One simple but effective technique is to increase the point value of exams throughout the term without increasing the length of the exam (three exams: 100, 120, and 140 points). It is worth more to do better later, so students receive the following messages: You do not have to be perfect at the outset, and practice has tangible value. It is likely that our students overestimate the modest

mathematical value of this scheme. We give a final examination (240 points) in which we attempt to provide an even representation of the term. One way I gauge improvement is to compare these two measures of cumulative performance: the average of the three exams given during the term and the average on the final. Using a simple spreadsheet, I subtract the final exam average (as a percentage) from the average of points for the three exams (as a percentage). By ranking the class according to this difference, I can locate the subset of my class whom I call the horizontal performers: those whose difference falls between +8 and –8 percent (a range I have designated, through experience, as being within a reasonable variance for identifying a significant difference).

I first assign my grade breaks to the horizontal performers. I think that students improve as a result of skill building, so I look to the group above +8 percent as a separate group for grading. On the other hand, I think that students decline for enough different reasons that I cannot say for sure that it is due to loss of skills. Once I have set my grades for the horizontal group, I sort back in the decliners and award them the grade based solely on their scores, as determined by the grid of grades already established. However, I think that Student A, who accumulates 300 course points by scoring 20/100, 50/120, 50/140 (33 percent), and then 180/240 (75 percent), is fundamentally different from Student B, who scores 50/100, 60/120, 70/140 (50 percent) and then 120/240 (50 percent). Student B is a horizontal performer whose grade would correspond to whatever "50 percent" ended up being that term, while I want to recognize that Student A has improved. To assign a grade for these incliners, I use the final exam papers, the placement of their averages, and all due fairness to students with consistent performances. In this case, Student A would definitely receive a higher grade than Student B in my course, but not the grade earned by someone who scored at the 75th percent level consistently during the term.

- *Use an absolute scale*. Setting an absolute scale means more than saying that 90 to 100 percent is an A grade. Our system depends on the fact that we give common examinations and fundamentally agree on course standards. These standards were determined empirically. By the third year of Structure and Reactivity, we had enough experience to be able to set guidelines for performance based on correlating numerical values with the rich and informative student work presented in their papers. This system would not be easy with multiple-choice examinations. We have set our examination standard high, and we are comfortable with the idea that

Siebert, Eleanor D. (Editor); McIntosh, William J. (Editor). College Pathways to the Science Education Standards

achievement above (or below) certain levels tells us about student performance (for more detail, see Coppola, Ege, and Lawton 1997).

• *Involve students in the process.* To grade examinations, we use graduate student instructors who are divided into groups. This technique has evolved over the years to a point where we can achieve a high level of consistency scoring over 1,000 exams with a staff of roughly forty graders in a four- or five-hour period. Since 1996, I have used a technique that attempts to demys-tify the grading process for my undergraduate students. During the grading session for the first examination, I look for two problems on which there are high variations in student responses. Before they are graded, I photocopy (or scan) the responses of four to six students. I then combine these into a handout with all identifiers to the originators removed. During class the next day, and prior to posting the exam key, I use the first twenty-five minutes to do an analysis of this handout. The total point values associated with the problems appear on the page. The students work in small groups to consider the answers to these problems and to create a fair grading scale, given the point values. This is, of course, exactly what I must have done prior to the grading session. After ten minutes, I call for the grading schemes and bring this discussion forward. Remarkably, the students will converge on the scheme that I created the previous evening within a point or two. This is an empowering experience. Even though the grading is already completed, my students get the sense that their thinking and my thinking converge. With the remaining time, I give the final grading scheme for these two problems and direct the groups to actually assign scores. Finally, I once again facilitate the feedback to the class by surveying the values assigned by members of the class to the work of their peers.

Making Sound Inferences



The inferences made from assessments about student achievement and opportunity to learn must be sound.

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

Source: National Resource Council. (1996). *National Science Education Standards*. Washington, DC: National Academy Press, 86.

Assessment Standard E appears to state the obvious: (1) do not stretch the data to support unwarranted conclusions, and (2) clearly state any assumptions on which inferences are based. Too often data are used to support conclusions that are very loosely linked to data and without an understanding of conditions of the study. For example, factors that must be known when interpreting data include sample size, control or reference groups used, and external factors that have been controlled (and those which have not been controlled). Assessment Standard E may be regarded as a summary statement designed to keep everyone—teachers of science, administrators, and those who make policy decisions that affect science education—alert to the need for objective interpretation of assessment data. This requires looking at assumptions underlying the study and the data collected to make certain that the inferences on student achievement and opportunity to learn are valid.

At the college and university level, sound inference from assessment is feasible if assessments are carefully designed by faculty and related to the outcomes of the course and program. At Alverno College over the past twenty-five years, faculty have identified clear overall outcomes for college education. With a strong emphasis on collaborative work, it has been natural for faculty to discuss the meaning and interpretation of this standard in a collaborative way. The first vignette below illustrates how Assessment Standard E is reflected in the process used to design assessments in courses at Alverno. In the second vignette, the assumptions on which class assessment are based are included as a list of criteria, referred to as a heuristic.

From the Field

Assessment as Student Learning

Leona Truchan, Department of Biology
George Gurria and Lauralee Guilbault, Department of Chemistry
Alverno College

This essay addresses student learning and development through the design and implementation of assessment; our focus is assessment designed for assisting individual students. At Alverno College, an office of research and evaluation focuses on program evaluation. Results from assessments that faculty have developed have been used by this office to identify program and curricular effectiveness (a summary of the findings of the office of research and evaluation appears in *Assessing and Validating the Outcomes of College* [Mentkowski and Loacker 1985]). Our interpretation of student assessment is driven by a clearly defined framework, evolved by the faculty and informed by our best practice (Alverno College Faculty 1994). Designing an instrument based on these components is essential to make certain that assessment is truly a learning opportunity for the student.

Siebert, Eleanor D. (Editor); McIntosh, William J. (Editor). College Pathways to the Science Education Standards

Siebert, Eleanor D. (Editor); McIntosh, William J. (Editor). College Pathways to the Science Education Standards.
Arlington, VA, USA: National Science Teachers Association, 2001. p 87.
http://site.ebrary.com/lib/brandeis/Doc?id=10240805&ppg=106

Copyright © 2001. National Science Teachers Association. All rights reserved.
May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.