

Features

Approaches to Biology Teaching and Learning: From Assays to Assessments—On Collecting Evidence in Science Teaching

Kimberly Tanner^{*†} and Deborah Allen[‡]

^{*}San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, and [‡]Department of Biological Sciences, University of Delaware, Newark, DE 19716

BRINGING THE CULTURE OF EVIDENCE TO THE BIOLOGY CLASSROOM

As scientists, we are accustomed to operating in a professional culture of evidence. Evidence is employed in the public discourse of science to support new ideas within a field and refute old ones. We can refer to this form of evidence—thorough, detailed, extensively reproduced and analyzed—as a *summative* form of scientific evidence. It is this summative form of evidence that is presented in conference proceedings, in journal publications, and, eventually, in the more static context of books. Yet evidence is also collected and used in more local and iterative ways in the daily life of the scientific laboratory. Results from preliminary investigations guide the design of larger-scale studies, and more often than not entire new lines of inquiry have their origins in unanticipated results that emerge from an experiment designed to address a wholly different question. Who among us has not performed the exploratory experiment to guide our ideas and explore areas of interest without the commitment of all the controls and multiple trials that we would require of a more mature experiment? We can refer to this form of evidence—exploratory, preliminary, informative, and instructive for future experiments—as a *formative* type of scientific evidence, one that is much less publicly acknowledged, although it may be shared and discussed.

Although both formative and summative evidence is the currency of knowledge and decision-making for scientists in the laboratory, evidence of any systematic sort has played a comparatively minimal role for scientists in their teaching practice. In science classrooms, evidence is often employed only *summatively*, in the assignment of grades for an exam or course and as a necessary means to inform students of a final judgment of their learning. More rarely evidence in science teaching and learning is used *formatively*, in gauging student understanding, identifying confusions, and guiding instruc-

tion on a daily basis. Of all the arenas of learning in schools and universities, one would expect the sciences to embrace fully the culture of evidence, both formative and summative, in the practice of teaching. Yet this is often not the case. How can we as scientists not be driven by such questions as: What do we want our students to learn? How do our students think about biology? and How can we adapt our teaching practices to better align student learning with our goals for student learning? Formative evidence in science teaching, in the form of classroom assessments, can play a key role in allowing scientists to pursue these questions and to bring a culture of evidence to the teaching and learning of science. Below we provide an overview of classroom assessment, as well as descriptions of several key resources that provide additional background information, assessment tools, and analysis techniques for embarking on new ventures in classroom assessment.

WHAT IS CLASSROOM ASSESSMENT? WHAT IS IT NOT?

“Formative assessment,” “student-centered assessment,” “embedded assessment,” “learner-centered assessment,” and “classroom assessment” (the term we use hereafter) are all monikers that can be used to describe the type of assessment that gives insight into the understanding of the learner, informs teaching practice, and is embedded in the culture of the classroom. In their *Classroom Assessment Techniques: A Handbook for College Teachers*, Thomas Angelo and K. Patricia Cross (1993) outline seven assumptions about what *classroom assessment* is, providing greater definition than simple names can convey (see Table 1). Importantly, Angelo and Cross emphasize that central to understanding the role of classroom assessment is the acknowledgment of the interconnectedness of learning and teaching. Effective teaching is fundamentally about student learning. Though seemingly simple, bridging the divide between teacher and student, and between what is taught and what is understood, can be very difficult. To accomplish deep

DOI: 10.1187/cbe.04-03-0037

[†]Corresponding author. E-mail address: kdtanner@sfsu.edu.

Table 1. The seven basic assumptions of classroom assessment: Adapted from Angelo and Cross (1993)

1. The quality of student learning is directly, although not exclusively, related to the quality of teaching. Therefore, one of the most promising ways to improve learning is to improve teaching.
2. To improve their effectiveness, teachers need first to make their goals and objectives explicit and then to get specific, comprehensible feedback on the extent to which they are achieving these goals and objectives.
3. To improve their learning, students need to receive appropriate and focused feedback early and often; they also need to learn how to assess their own learning.
4. The type of assessment most likely to improve teaching and learning is that conducted by faculty to answer questions they themselves have formulated in response to issues or problems in their own teaching.
5. Systematic inquiry and intellectual challenge are powerful sources of motivation, growth, and renewal for college teachers, and Classroom Assessment can provide such challenge.
6. Classroom Assessment does not require specialized training; it can be carried out by dedicated teachers from all disciplines.
7. By collaborating with colleagues and actively involving students in Classroom Assessment efforts, faculty (and students) enhance learning and personal satisfaction.

understanding in a discipline, educators must move beyond the traditional practices of *telling as teaching* and *memorizing as learning*. Classroom assessment is a key tool in connecting learning to teaching and identifying that which is *not* being understood by students and what alternative conceptions or misconceptions students hold about the natural world. The practitioner of classroom assessment is anyone who is an instructor, and as such, classroom assessment does not require specialized training but, rather, is within the domain of anyone guiding learning by students. Angelo and Cross complete their list of assumptions by pointing out that classroom assessment is a collaborative effort among teachers and students, in which students are actively engaged in reflecting on their own understanding.

As Paul Black, physicist and assessment specialist, has eloquently expressed time and again, assessment can serve at least three major purposes: accountability, certification, and learning (Black and William, 1998; Atkin, 2002). Assessments in the service of *accountability*, such as the National Assessment of Educational Progress (NAEP) and the Third International Mathematics and Science Study (TIMSS), often involve large-scale, multisite testing efforts that are intended to inform policy and drive reform. Assessments in the service of *certification*, such as the SAT, the ACT, and the National Medical Board Exam, to name but a few, are examinations that determine educational eligibility or professional licensure. Classroom assessment, however, is neither about accountability nor certification, but rather about assessment in the service of *learning*. It is perhaps important to articulate further what classroom assessment is not.

- *Classroom assessment is NOT about proving success.* It is wonderful when the results of an assessment show that your students are really “getting it”! However, more often than not, assessments yield important insights into what students are NOT getting and HOW and perhaps WHY they are not getting it.
- *Classroom assessment is NOT done for accountability to outside stakeholders.* Classroom assessment is clearly focused in the realm of the teacher and the learner, within the relatively intimate and unique setting that is every individual classroom. The outcomes of classroom assessment should be de-

signed to be useful to both the instructor and the students, not to external stakeholders.

- *Classroom assessment is NOT specifically about grading.* Although assessment may be linked with grades, grading in the traditional sense of the numeric labeling of the performance of a student is not the primary goal of classroom assessment.
- *Classroom assessment is NOT clean, neat, and perfectly orderly.* Classroom assessment, by its nature of exploring student thinking, can be messy, can involve several iterations, and is expected to produce different outcomes with different students.
- *Classroom assessment and its methodologies are NOT identical to scientific research and its methodologies.* Often scientists are hindered from conducting classroom assessments because of an expectation that any evidence collected in the classroom must resemble evidence that they collect in their laboratories. It is true not only that the nature of evidence differs in the classroom and laboratory, but that evidence differs widely across the span of scientific disciplines.

As shown in Table 2, while scientific research and classroom assessment do have commonalities in their reliance on evidence and their generation of new knowledge, they differ in their goals, subjects, and methodologies. By virtue of being grounded in a particular instructional setting, classroom assessment is highly local and not necessarily generalizable. That said, data emerging from classroom assessments can nucleate more extensive and systematic lines of inquiry and lead to classroom-based research, termed “action research,” an ongoing process of systematic, self-study in which individual instructors examine their own students’ learning in detail as an evidence base from which to improve their own teaching practice (Altrichter *et al.*, 1993; Loucks-Horsley *et al.*, 1998).

THE ITERATIVE NATURE OF CLASSROOM ASSESSMENT

As scientists embark on new ventures in classroom assessment, it is important to recognize the iterative nature of the process (see Figure 1). Classroom assessments are not an end in and of themselves but, rather, support a process of reflection on student understanding and teaching practice.

Table 2. Comparing and contrasting classroom assessment and scientific research

| | Classroom Assessment | Scientific Research | Commonalities |
|---------------------------|--|---|--|
| Goal | To understand student learning and inform teaching practice | To understand the natural world | To increase knowledge base and general understanding |
| Subject | Student and teacher interactions studied in a particular classroom | Commonly observed natural phenomena studied across multiple laboratories | Focus on investigation of specific questions about the subject |
| Methodology | Emphasizes understanding student learning, which is unique from classroom to classroom | Emphasizes controlling variables and reproducing results, which should be similar from laboratory to laboratory | Emphasizes systematic collection, analysis, and interpretation of evidence |
| Nature of Evidence | Primarily focused on formative evidence | Primarily focused on summative evidence | Evidence drives iterative process, generating new questions |
| Outcomes | Expands knowledge base of individual instructor about teaching effectiveness | Expands knowledge base of scientific community about scientific phenomena | Contributes to an evidence-based body of knowledge |

Classroom Assessment Is About Asking Questions About Student Learning

The main goal of classroom assessment is to better understand the relationship between what students learn and what we think we are teaching them. As such, classroom assessments are simply methods to aid instructors in answering questions about what and how our students are learning. What do you wonder about what your students are learning? How do you access what your students already know? What misconceptions do they bring with them to the classroom?

Methods for Collecting Classroom Assessment Data Are Guided by the Questions

Just as multiple assays and experimental approaches are available for discovering new knowledge in the laboratory, so are multiple assessment methodologies available for investigating student understanding. Debates have long occurred in the field of educational assessment about the relative richness, validity, and appropriateness of different assessment methodologies, in particular, quantitative versus qualitative instruments (Sundberg, 2002). It is important to realize is that there is no one right approach to classroom assessment. Rather, the choice of assessment methodology should be based on what type of evidence will provide insight into your question about student learning. For example, concept maps are an excellent

tool for understanding the breadth of knowledge and connections among concepts held by students on a given topic. However, concept maps are much less appropriate and less effective in assessing students' ability to analyze and interpret experimental data. More appropriate to assess students' skills in evaluating data would be a performance-based assessment in which students are presented with actual data for analysis, collected either by themselves or by the instructor or from scientific research papers by other scientists.

Analysis of Classroom Assessments Leads to Instructional Choices and New Questions

At their most effective, classroom assessments will inform future instructional choices. Classroom assessments can yield insight into what students already know and what misconceptions they have. These insights can in turn guide the relative emphasis, time spent, and teaching strategies used in building student knowledge. While building that understanding, classroom assessments can continually play a role in probing student ideas, gauging whether misconceptions are being resolved or persisting, and detecting unanticipated conceptual challenges. Unsurprisingly, analysis of classroom assessment data often leads to more questions, not unlike experimental results in the laboratory. Thus, when embarking on classroom assessment, instructors should expect to find themselves engaged in a cyclical venture (see Figure 1).

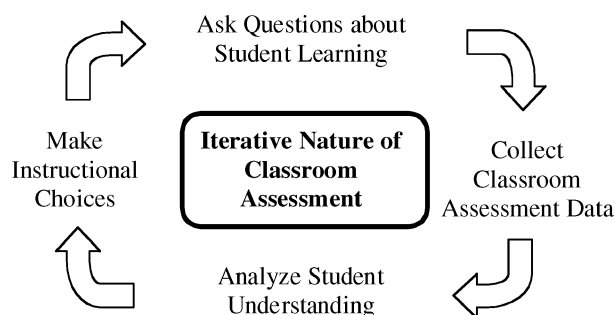


Figure 1. The Iterative Nature of Classroom Assessment.

RESOURCES ON CLASSROOM ASSESSMENT TO GUIDE THE WAY

Although interest in classroom assessment may be high, oft-heard statements from colleagues include, "But where would I start with assessment in my classroom?" and "I don't know anything about assessment." The following resources can provide scientists entry points into the literature on classroom assessment and are widely regarded as rich resources for instructors embarking on their own action research or classroom assessment projects.

An Introduction to College-Level Classroom Assessment

***Classroom Assessment Techniques: A Handbook for College Teachers*, by Thomas A. Angelo and K. Patricia Cross (1993).**

This compendium by Angelo and Cross is currently one of the most comprehensive guide to classroom assessment available for college and university instructors. It provides easy entry into the philosophy of formative classroom assessment, as well as describes methodologies available to gather evidence of student learning in the classroom. Making few assumptions about the background of the reader, the guide begins with an overview entitled "Getting Started in Classroom Assessment," in which the authors make explicit their seven basic assumptions (see Table 1). In addition, the reader is prompted to conduct a self-evaluation, "The Teaching Goals Inventory," to emphasize the centrality of instructional goals in designing classroom assessments. Angelo and Cross then present over 50 different classroom assessment techniques (CATs), derived from the education research literature, their own instructional practice, and the repertoires of other faculty. These techniques are then organized into three sections, identifying tools most appropriate for 1) assessing course-related knowledge and skills, 2) assessing learner attitudes, values, and self-awareness, and 3) assessing learner reactions to instruction. Although some techniques presented are more widely known, such as Concept Mapping, Minute Papers, and The Muddiest Point, many will be novel even to those with extensive experience in classroom assessment, including techniques entitled Defining Features Matrix, Approximate Analogies, and Directed Paraphrasing. Although no single assessment tool is delved into deeply—for example, concept mapping occupies a mere four pages—each is accompanied by an example, a step-by-step procedure, the pros and cons of the particular technique, suggested situations for using the technique, and an alignment with particular teaching goals for which the technique is most appropriate. At first glance, scientists may note that the content areas represented by Angelo and Cross include disciplines as diverse as nursing, economics, anthropology, music, literature, and foreign language. That said, the examples offered, whether or not in a science field, are generally detailed enough to serve as models for the development of a similar classroom assessment in one's own field.

On Classroom Assessment in College-Level Science

***The Field-Tested Learning Assessment Guide* (www.flaguide.org).** Developed by the College Level One Team at the National Institute for Science Education (NISE) (2003), based at the University of Wisconsin—Madison, the Field-Tested Learning Assessment Guide (FLAG) is an excellent and accessible starting point for instructors who wish to expand their knowledge of classroom dynamics and access a variety of assessment tools and resources. The FLAG website gathers in one place assessment techniques specifically designed for courses in science, mathematics, engineering, and technology. Providing a wealth of well-referenced resources, FLAG is organized into five areas: 1) A Primer on Assessment, 2) Teaching Goals, 3) Classroom Assessment Techniques, 4) Specific Assessment Tools, and 5) Resources in Assessment. Following the general introduction to assessment, the CATs

section provides an introduction to general methods of assessment such as attitude surveys, interviews, weekly reports, portfolios, ConcepTests, Minute Papers, and Concept Mapping. The description of each CAT presented is written by a college or university instructor who has implemented the technique, and each CAT underwent a peer review process. For example, three chemistry instructors from the University of Wisconsin—Madison describe their implementation of the ConcepTest assessment tool, a technique originally developed for use in large-class physics lectures by Harvard University professor Eric Mazur (1996). In employing a ConcepTest,

the instructor presents one or more questions during class involving key concepts, along with several possible answers. Students in the class indicate by, for example, a show of hands, which answer they think is correct. If most of the class has not identified the correct answer, students are given a short time in lecture to try to persuade their neighbor(s) that their answer is correct. The question is asked a second time by the instructor to gauge class mastery. Many variations on this general CAT exist. A video clip illustrating the method is part of this CAT description.

Although the probing question detailed on the website is specific to chemistry, the detailed description of the methodology provides an excellent model for developing ConcepTests as classroom assessments in large classrooms in any content area. The Tools section comprises a database of specific assessment instruments that can be sorted by either discipline or type of methodology. For biologists searching the database, it will become immediately apparent that life scientists are in need of more classroom assessment instruments, perhaps similar to those that have been developed in chemistry and physics (Klymkowsky *et al.*, 2003). Finally, for the reader who wishes to pursue a particular classroom assessment topic in more depth, the Resources section includes information about other assessment websites, assessment experts in your area of the country, and an annotated bibliography of books on assessment, a limited number of relevant assessment articles, and links to over 30 science education journals. While FLAG is generally congruent with the research and publications of Angelo and Cross (e.g., Angelo and Cross, 1993; Cross and Steadman, 1996), its strength lies in the fact that it is specific to the content areas of science, mathematics, and engineering. In addition, since FLAG is archived as a Web site, its online accessibility is an asset, though we note that the project is no longer in active development; consequently, the materials available at FLAG are likely to become increasingly outdated.

Resources on Classroom Assessment Rubrics and Analysis

***Learner-Centered Assessment on College Campuses*, by Mary E. Huba and Jann E. Freed (2000).** In this book, subtitled *Shifting the Focus from Teaching to Learning*, Huba and Freed have crafted a detailed, thoughtful, and thorough introduction to employing classroom assessment in the service of student learning. Practicing what they preach, the authors carefully embed throughout the book frequent self-assessment text boxes with questions that prompt the reader to consider

prior knowledge and experiences, as well as to strategize about implementation of assessment tools and predict potential outcomes. The forte of this particular resource, though, lies specifically in two chapters. Both Chapter 5, "Using Feedback to Improve Student Learning," and Chapter 6, "Using Rubrics to Provide Feedback to Students," provide guidance for the reader on what to do with classroom assessment data once collected, a topic to which the above resources only allude. Specifically, in Chapter 6, Huba and Freed delve deeply into the topic of rubrics, tools that make explicit and public an instructor's criteria for evaluating and scoring classroom assessment data. The authors present three sample rubrics, deconstruct these rubrics, and emerge with a very practical guide for developing useful rubrics for classroom assessments. Once classroom assessment data have been collected and analyzed, the authors go further to discuss approaches to sharing insights from assessments with students. Both their guidelines for effective feedback discussions and their questioning techniques in support of these discussions are unique and useful tools for closing the loop and taking the results of classroom assessments back to students.

Effective Grading: A Tool for Learning and Assessment, by Barbara E. Walvoord and Virginia Johnson Anderson (1998). Published in 1998, this resource addresses what for many is a continuing conundrum, namely, how to connect classroom assessment with traditional demands for assigning students grades. Similarly to Huba and Freed, these authors outline strategies for establishing criteria and standards for grading and detail the design of "primary trait analysis scales," tools for analyzing assessment data similar to rubrics. Unlike Huba and Freed, however, these authors pursue more practical aspects of the intersection between grading and classroom assessment by addressing topics such as "managing the grading process," "calculating course grades," and "making grading more time efficient." The extent to which classroom assessments and grading overlap is a worthy topic in and of itself, and the curious reader will be rewarded by exploring the ideas presented.

Resources on K-12 Science Classroom Assessment

Everyday Assessment in the Science Classroom, Edited by J. Myron Atkin and Janet E. Coffey (2003). This collection of essays published by the National Science Teachers Association considers classroom assessment in K-12 science classrooms. While covering some of the same topics as the resource guides described above, this book explores topics that are not addressed in the college-level guides. Most notably, in his essay on "Assessment of Inquiry," Richard Duschl argues for the importance of listening to student discussion, argument, and debate as a key method of collecting evidence on student understanding of scientific inquiry. Similarly, the importance of scientific discourse, questioning strategies, and teacher listening is highlighted in the chapter entitled "Using Questioning to Assess and Foster Student Thinking," by Jim Minstrell and Emily van Zee.

Assessment and the National Science Education Standards, Edited by J. Myron Atkin, Paul Black, and Janet E. Coffey (2001). Produced by two of the same editors as *Everyday Assessment in the Science Classroom* and published by the National

Research Council, this book was published as a companion volume to the *National Science Education Standards* (NRC, 1996). Compiled as an overview intended for K-12 teachers, it is an interesting cousin to the aforementioned college-level guides. Most informative, and unique among all the resources listed here, are the specific examples describing what classroom assessment looks like in a variety of K-12 classrooms. These examples are predominantly drawn from classroom observations collected by science education researchers and provide a unique view of what daily classroom assessment really looks like, a view that is not widely available for college- and university-level classrooms.

CLASSROOM ASSESSMENT BEYOND THE CLASSROOM

Articulated as the seventh and final assumption about classroom assessment (see Table 1) is the potential role of collaboration in the process. There is enormous potential in collaborative faculty groups engaging in the development and examination of science assessments, whether across sections of a single course, across different courses in a discipline, or even across different disciplines in the fields of science and mathematics. Such steps toward collaboration in classroom assessment could begin to establish the use of evidence in teaching as a cultural norm in the sciences. In addition, discussion of classroom assessments with colleagues outside of one's own classroom has the potential to nucleate scholarly efforts in the realm of college science teaching. Classroom assessments, while initiated for the betterment of teaching and learning, can produce unanticipated results and insights of interest to a larger audience. Taken to its logical end, classroom assessments used formatively in science teaching can mature into classroom research in a more summative form. As Patricia Cross writes in *Classroom Research: Implementing the Scholarship of Teaching*, "Classroom assessment typically answers questions about what students are learning and how well, but it often raises questions about *how* students learn. Those questions lead teachers to Classroom Research" (Cross and Steadman, 1996). It is at this point that classroom assessments may also play a role outside the classroom in providing evidence for the effectiveness of instructional strategies and promoting the scholarship of teaching (Cross and Steadman, 1996; Sundberg, 2002). In this way, endeavors in classroom assessment may lead you to the Instructions for Authors page of this very journal.

REFERENCES

- Altrichter, H., Posch, P., and Somekh, B. (1993). *Teachers Investigate Their Work: An Introduction to the Methods of Action Research*, London: Routledge.
- Angelo, T.A., and Cross, K.P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers*, San Francisco, CA: Jossey-Bass.
- Atkin, J.M., and Coffey, J.E., eds. (2003). *Everyday Assessment in the Science Classroom*, National Science Teachers Association Press.
- Atkin, J.M., Black, P., and Coffey, J.E., eds. (2001). *Assessment and the National Science Education Standards*, Washington, DC: Center for Education, National Research Council.

- Black, P., and Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan* 80(2), 139–148.
- Cross, P.K., and Steadman, M.H. (1996). *Classroom Research: Implementing the Scholarship of Teaching*, San Francisco, CA: Jossey–Bass.
- Huba, M.E., and Freed, J.E. (2000), *Learner-Centered Assessment on College Campuses*, Needham Heights, MA: Allyn and Bacon.
- Klymkowsky, M.W., Garvin-Doxas, K., and Zeilik, M. (2003). Bioliteracy and teaching efficacy: What biologists can learn from physicists. *Cell Biol. Educ.* 2, 155–161.
- Loucks-Horsley, S., Hewson, P., Love, N., and Stiles, K. (1998). *Designing Professional Development for Teachers of Science and Mathematics*, Thousand Oaks, CA: Corwin Press/National Institute for Science Education.
- Mazur E. (1996). *Peer Instruction: A User’s Manual*, Upper Saddle River, NJ: Prentice Hall, 253.
- National Institute for Science Education. (2003). *Field-Tested Learning Assessment Guide*. www.flaguide.org.
- National Research Council. (1996). *National Science Education Standards*, Washington, DC: National Academy Press.
- Sundberg, M. (2002). Assessing student learning. *Cell Biol. Educ.* 1, 11–15.
- Walvoord, B.E., and Anderson, V.J. (1998). *Effective Grading: A Tool for Learning and Assessment*, San Francisco, CA: Jossey–Bass.