JOURNAL OF CHEMICAL EDUCATION

EDITORIAL

pubs.acs.org/jchemeduc

Assessment Data and Decision Making in Teaching

Thomas Holme*

Department of Chemistry, Iowa State University, Ames, Iowa 50011, United States

ABSTRACT: Finding avenues to collect and use data in our teaching parallel to the way we use data in research may be an important way to improve educational outcomes.

KEYWORDS: General Public, Public Understanding/Outreach, Testing/Assessment

Teaching is at once inherently personal and inescapably corporate. While it may not be commonly stated in this specific way, this fundamental dichotomy plays the paramount role in the management of teaching in higher education. How well students learn in a given class (or given educational institution) clearly matters to many people. In nearly all college classes, however, the decisions that most profoundly affect the learning that occurs are made by a sole individual—the instructor. Increasingly, the aspirations of the wider interests of education are being articulated by demands for assessment. A search on the terms "assessment" and "higher education" in the ERIC (Education Resource Information Center) database[1] reveals that over 1500 articles in 2010 include both of these search terms. How does this trend actually affect chemistry teachers? Should we be paying more attention to decisions about assessment and decisions made with data from assessment?

One key challenge for many chemistry teachers is gaining facility with educational measurements that comprise assessment efforts. As chemists, our field has advanced largely on the basis of precise measurements, even for very complex systems. Even though most academics probably recognize that tests or other forms of assessing learning are measurements, to chemists these measurements seem different enough that they tend to elicit different critical judgments. In the laboratory, when a measurement fails to provide needed data (or yields data contrary to the model being tested), considerable effort is often spent to redesign the experiment. In the classroom, however, when a test shows low student learning, there is a tendency to fault the system being measured (the group of students) more than the design of the measurement—the test. It may well be that something has impeded learning under such circumstances, yet a more scientific response would be to try to better understand characteristics of (i) the test; (ii) the students; and (iii) the interaction between the students and the test, or many other possible variables, including the nature of the instruction. In other words, the data would suggest efforts to redesign the system.

There is, of course, one major challenge related to this analogy. Instruments used to measure student learning are often not as well characterized as laboratory instruments in the physical sciences. One advantage to chemists of an institution like the ACS Exams Institute[2] is that it can often provide a benchmark for other forms of assessment.

In this *Journal*, we have described the ACS exam development process,[3] and recently we are providing more detailed analyses for exams that may be useful for research or evaluation purposes.[4] Norm-referenced, multiple-choice exams are not capable of measuring all aspects of learning; however, the availability of tests with established validity and reliability[5] can assist in the development of other measures. Educational innovations remain valuable, yet if those innovations are judged solely (or largely) on the basis of surveys that show "students like it", we are not subscribing to a standard any research scientist would expect.

Measurement of learning will always be challenging, and we cannot forego educational innovation simply because the likely gains will not be readily measured. Nonetheless, if we pursue educational change without simultaneously pursuing better measures of what students have learned, we will be starved for data. Assessment needs to be an equal partner with other forms of scholarship in education if we can ever hope to take a scientific tack in our efforts to improve learning.

AUTHOR INFORMATION

Corresponding Author

*E-mail: taholme@iastate.edu.

REFERENCES

(1) The ERIC Database is the largest education database in the world; many academic libraries subscribe to it. http://www.eric.ed.gov/ (accessed Jun 2011).

(2) American Chemical Society Division of Chemical Education Examinations Institute Home Page. http://chemexams.chem.iastate.edu/ (accessed Jun 2011).

(3) Holme, T. Assessment and Quality Control in Chemistry Education. *J. Chem. Educ.* **2003**, *80*, 594–598.

(4) Holme, T. Assessing Conceptual and Algorithmic Knowledge in General Chemistry with ACS Exams. *J. Chem. Educ.* **2011**, *88*, in press.

(5) Validity and reliability are specific quality estimates for assessment instruments; see, for example: Crocker, L. M.; Algina, J. *Introduction to Classical and Modern Test Theory*; Harcourt Brace: Fort Worth, TX, 1986.

Published: June 15, 2011

dx.doi.org/10.1021/ed200350w | *J. Chem. Educ.* 2011, 88, 1017–1017