

ASSESSMENT NOT EVALUATION IS THE KEY TO LEARNING

A recent summary of research into how people learn identified that effective learning environments need to be knowledge centered, learner centered, community centered, and assessment centered (1). Of these four, the least progress has been made in moving assessment to the center of the STEM (science, technology, engineering, and mathematics) classroom.

Content knowledge traditionally has been strong in STEM courses. Learner and community centered classrooms that engage students in the learning process and involve learning teams and learning communities are becoming more and more popular, and team problem-based learning has been a tradition in engineering, business, and medical programs. In contrast, only evaluation, and not assessment, commonly is found in STEM classrooms, even though assessment is essential to learning.

Evaluation is the process of measuring the quality of performance against standards to determine if the standards have been met. Assessment is the process of measuring and analyzing performance to provide timely feedback to improve future performance. Assessment differs from evaluation in that it provides opportunities for feedback, improvement, and revision; identifies strengths, areas for improvement, and ways to improve; and is non-judgmental, frequent, and congruent with the learning goals (2).

In most STEM college classrooms, feedback occurs relatively infrequently. Grades on quizzes, homework and examinations are perceived by students as evaluation intended to measure the result of learning and the level of accomplishment at a particular time. After receiving such grades, students typically move on to new topics and do not reflect on how to improve their understanding or skills in the graded topics or how to improve their learning process so they can improve their grades in future topics. Feedback is most valuable when it is accompanied by the opportunity to improve the process, and includes information on how this improvement can be accomplished.

The reporting of a single numerical grade is too imprecise for students to improve their learning process or for teachers to improve their teaching strategies. A single grade represents an average over several skill and knowledge items and fails to assess specific skill and knowledge components

Including a list of right and wrong answers with the grade also is insufficient to identify deficiencies in specific learning or problem solving processes because such a list provides little insight by itself. While it is possible to conduct an item analysis on examinations, probably no students and few faculty actually do this. Students have difficulty identifying the reasons they obtained the incorrect answer, and faculty find the effort needed to conduct and analyze an item analysis tedious and time consuming.

David M. Hanson
Professor of Chemistry & Chair
of the Learning Communities
Program
State University of New York -
Stony Brook

Assessment is an essential
component in the process to
improve learning and teaching.

This essay was prepared for the
2002 PKAL Roundtable on the
Future, *Assessment in the Service of
Student Learning*

ASSESSMENT NOT EVALUATION IS THE KEY TO LEARNING

Interviews or mentoring sessions by trained and knowledgeable staff can provide reliable data on learning and quality feedback to students, but such interviews have serious cost and implementation drawbacks that render them impractical as other than occasional learning assessment tools.

Too often the evaluations are not congruent with course goals. While essentially all teachers want their students to understand fundamental concepts and develop valuable skills, too many give exam and homework problems that can be solved by algorithms, pattern matching, and memorization. Students identify the course goals from the exams not from the philosophy of the instructor, behave accordingly, and never develop conceptual understanding and problem-solving skills.

In summary, current assessment practices are not producing an effective learning environment in STEM classrooms because they are actually summative evaluation not formative assessment, too infrequent, not precise, often incongruent with course goals, and inadequate in providing feedback to students, instructors, and organizations. These issues need to be addressed in any process that seeks to improve learning and teaching.

In addition, concern must be directed at the level of knowledge students need to succeed in courses, and assessment should be used to help students achieve at the highest levels.

According to constructivist learning theory, people construct their own understanding by integrating incoming information with information and experiences already in long-term memory. Students often do not appear to engage in this integration. Rather, they appear to rely most often on memorization, pattern matching, and algorithms to earn points rather than using the understanding they should have developed. Current evaluation practices do little if anything to discourage them from this strategy, and they manage to succeed by working with the lowest levels of knowledge.

At least four levels of knowledge should be considered in STEM courses, and assessment and evaluation instruments designed to measure performance at each of these four levels. These levels have been adapted from Bloom's taxonomy of educational objectives (3).

1. Information - characterized by memorization, able to recall and repeat pieces of information and identify information that is relevant.

2. Algorithmic Application - characterized by the ability to mimic, implement instructions, and to use memorized information in familiar contexts.

3. Comprehension/Conceptual - characterized by the ability to visualize, rephrase, change representations, make connections, and provide explanations.

4. Analysis/Synthesis - characterized by the ability to use material in new contexts (transference); to analyze problems; to identify the information, algorithms, and understanding needed to solve them; to synthesize these components into a solution; and to evaluate the quality of the solution.

Formulating questions and analyzing responses at the information and algorithmic levels generally is straightforward as such questions are very directed and typically require a single response that often is generated in a single step. Differentiating learning at the comprehension/ conceptual level has been discussed in the literature, (4,5,6) and conceptual understanding can be assessed through single or coupled questions that have certain characteristics. Such questions have been used previously in the physics Force Concept Inventory, (7, 8) Mazur's ConcepTests, (9) and the ACS General Chemistry Conceptual Examination (10). Such concept questions tend to have one of ten characteristics.

- ◆ Require qualitative answers about situations described by mathematical equations.
- ◆ Present a situation to analyze and require a prediction or a qualitative conclusion.
- ◆ Ask to change a representation, interpret it, or conclude from it. The representation might be nanoscopic, macroscopic, pictorial, graphical, mathematical, symbolic, tabular, or verbal.
- ◆ Use equations with proportions rather than numbers so plug and chug isn't possible.
- ◆ Take a concept and require that it be used in different contexts to reach conclusions.



ASSESSMENT NOT EVALUATION IS THE KEY TO LEARNING

- ♦ Ask a question plus a follow-up asking for a reason.
- ♦ Require matching a statement and a reason.
- ♦ Require completing a statement to make it valid and/or to explain why it is valid.
- ♦ Ask a question with a response that requires making connections between two or more concepts.
- ♦ Ask that correct or incorrect statements to be identified.

Questions at the analysis/synthesis level are straightforward to identify and develop. Typically these are called problems. Some characteristics of questions that represent problems for students are given below.

- ♦ Require transference of prior learning to new contexts.
- ♦ Require synthesis of material from different contexts.
- ♦ Require analysis to identify information provided, information needed, algorithms and concepts required to connect these, followed by synthesis into a solution.
- ♦ Require developing a model, making estimates, making assumptions.
- ♦ Are missing information.
- ♦ Have extraneous information.
- ♦ Have multiple parts.
- ♦ Lack clues regarding the route to the solution and the relevant concepts.

The bottom line is that assessment is a key instructional tool that can be used to improve the teaching/learning process, and this tool also can be used to raise the level of knowledge performance by STEM course students.

References

1. J.D. Bransford, A.L. Brown and R.R. Cocking, eds. *How People Learn*. 1999, National Academy Press: Washington, D,C.
2. P.E. Parker, P.D. Fleming, S. Beyerlein, D. Apple and K. Krumsieg. *Differentiating Assessment from Evaluation as Continuous Improvement Tools*. In *Thirty-first ASEE/IEEE Frontiers in Education Conference*. 2001. Reno, NV.
3. B.S. Bloom, M.D. Engelhart, E.J. Furst, W.H. Hill and D.R. Krathwohl, *Taxonomy of Educational Objectives: Classification of Educational Goals, I. Cognitive Domain*. 1956, New York: David McKay Company.
4. C.W. Bowen, *Item Design Considerations for Computer-Based Testing of Student Learning in Chemistry*. *J. Chem. Ed.*, 1998. 75: p. 1172-1175.
5. C.W. Bowen and D.M. Bunce, *Testing for Conceptual Understanding in General Chemistry*. *Chemical Educator*, 1997. 2: p. 1430-1471.
6. K.J. Smith and P.A. Metz, *Evaluating Student Understanding of Solution Chemistry Through Microscopic Representations*. *J. Chem. Ed.*, 1996. 73: p. 233-235.
7. D. Hestenes, M. Wells and G. Swackhamer, Force Concept Inventory. *Phys. Teach.*, 1992. 30: p. 141-157.
8. D. Hestenes and I. Halloun, *Phys. Teach.*, 1995. 33: p. 502, 504.
9. E. Mazur, *Peer Instruction*. 1997, Upper Saddle River, NJ: Prentice Hall.
10. I.D. Eubanks, *ACS Division of Chemical Education: General*

Chemistry (Conceptual). 1996, Clemson University: Clemson, SC. ■