Third Annual National Institute for Science Education
Indicators of Success in Postsecondary Science, Mathematics, Engineering and Technical Education: Shapes of the Future.

Session One - Assessment in the Classroom

The integrality of assessment — David B. Porter, US Air Force Academy

Before addressing the questions posed for this session, it is important for you to understand a little about the context and perspective in which my ideas developed. Although all institutions of higher education are unique, federal military academies are, in some respects, in a category of their own. Many of the fundamentals of teaching, learning, and motivation, however, are likely to be similar to those encountered at other colleges and universities. The U.S. Air Force Academy is a four-year undergraduate university. Its mission is *to develop and inspire air and space leaders with vision for tomorrow*. Academy cadets (students) are competitively selected and do not pay tuition. They incur a commitment to serve as commissioned officers in the Air Force for five years after graduation. The size of the Cadet Wing (student body) is about 4,000. About 14% are women and 18% represent racial minorities. As a group, cadets are bright (average SAT scores are typically above 1300), athletic (more than 80% earned letters in varsity sports in high school) and tend to share conservative social and political perspectives.

The academic faculty of 530 is about 20% civilians and 80% military officers assigned to 19 academic departments in four academic divisions: Basic Sciences, Engineering, Social Sciences and Humanities. Since an assignment to the Air Force Academy is considered *a special duty* for Air Force officers, most military faculty members serve only a single three-year tour. As a result, the average teaching experience among faculty is only 2 years, and annual turnover in the academic departments often is 30%. Recent hiring of full-time civilian faculty has increased the proportion of faculty with doctorate degrees to about 50%. Class sizes are usually less than 20 students, and an average teaching load for junior faculty is 4 sections (12 semester hrs). Teaching and learning are emphasized across the faculty. Several years ago, governmental emphasis on *Quality* increased the focus on assessment as a way to enhance institutional effectiveness.

***Student learning** - How can faculty best understand what undergraduates are learning, retaining, and using in future contexts?*

Few engineers would entertain the notion of building a bridge contrary to the laws of physics. Although somewhat less precise, the relationships that gird the mental world of thinking and learning are just as potent as those that constrain the physical world. Faculty who design and deliver curricula without understanding the principles of human learning are likely to waste their own time and create classes that harm students.

The process of learning is susceptible to the same scientific method used in the physical sciences. A coherent model of human learning is a necessary starting point for such inquiry. Learning has adaptive significance; it is a natural phenomenon for humans; it enhances our individual and collective chances of survival (Gould, 1981). Learning assumes that certain experiences and activities enhance the capacity of the individual to deal with environmental challenges. Response quality (i.e., performance) can be influenced by many things (viz., knowledge, skills and attitudes). Changes in any of these components can affect both performance and learning. This suggests that the *question of student learning* might be more accurately considered as several questions: *What is known that wasn't known before? What can be done that couldn't be done before?* And *what is the effect on student attitudes?* (Porter, 1991).

The relationships between these components of learning are even more important than the components themselves. For example, the type of mental activity involved in study (i.e., a skill) is a much better predictor of retention than is duration of exposure (Craik & Lockhart, 1973; Craik & Tulving, 1975). *Elaborative rehearsal* typically involves activities such as reflection, comparison, argument and conclusion. *Maintenance rehearsal* is the act of simply repeating a phrase, formula, or particular "fact" verbatim. Elaborative rehearsal usually yields two to three times greater retention than maintenance rehearsal. Ironically, the most common student study strategy is to "bear down" and rely heavily on maintenance rehearsal, especially in those subjects students find most challenging or distasteful (i.e., science and math classes).

Another effective but under-employed strategy is *visualization*. Although there are individual differences, material that has been visualized is about twice as likely to be recalled as material that has been verbalized for an equal amount of time (Atkinson & Raugh, 1975). In fact, there is strong evidence that the greatest retention is likely to occur when presentations and activities involve both visual and verbal processing (Pavio, 1971). One final example of how process affects knowledge retention is known as the *self-referent effect* (Rogers, Kuiper, & Kirker, 1977). Students remember best what they care about most and what connects to them most personally. In a typical experiment, subjects are asked to rate one list of adjectives on a 5-point scale as to how positive or negative the words are. Another list of adjectives is also rated on a 5-point scale as to how well the words describe the subject personally. The usual result is that subjects recall three times as many adjectives that they rated in relationship to themselves.

Elaboration, visualization, and self-reference are not innate study techniques. In fact, many students (and faculty) steeped in an academic version of *the Protestant Work Ethic* (viz., *no pain, no gain*) assume such techniques are simply frivolous diversions. To enhance learning, these techniques must be presented, advocated, practiced, and actively incorporated into lessons and courses. This requires cooperation and trust between students and their teacher. To the extent students are fearful of failing, see the teacher as an adversary, or are engaged in direct interpersonal competition, the necessary collegial classroom climate is unlikely to develop (Kohn, 1986; Glasser, 1990; Palmer, 1997). Recognizing the importance of student attitudes and *driving out fear* are often prerequisites of pedagogical progress and development. Educational success is often contingent upon students seeing faculty as allies in their battle against a common enemy: *ignorance.*

However, all these techniques still do not directly address the question of "how" faculty can best understand student learning. While mastery of techniques may be necessary, no level of mastery is sufficient to assure insight. After several years of working with good people, who strongly desire to become effective educators, I'm convinced the single most critical variable is also one of the least tangible: *authenticity.* Teachers must be able to be themselves; "*who they are*" and "*what really matters to them*" is what creates the classroom climate and provides the crucible in which learning might occur. Students need a safe place for substantive conversations to take place before they will share the secrets of their private perceptions and assumptions. Within a supportive context, I've found a single three-word phrase to be very helpful in drawing out the information I need to understand what students are learning. These three words are: "*Help me understand.*" However, as a mere technique even these "magic words" are likely to be ineffective if the teacher doesn't really mean them or hasn't yet convinced the students of this. The words must "fit" the teacher and also the rest of the course; they must be "authentic".

Students report the greatest learning when faculty emphasize all three types of outcomes (knowledge, skills and attitudes). A study of 115 Air Force Academy faculty members suggested teachers who balance emphasis on students' knowledge, skills, and attitudes increase students' subsequent perceptions of their own learning. In fact, this effect was larger than the influence of teaching experience, teacher temperament, and degree-level combined (Porter and Benson, 1995).

As part of the Air Force Academy's recent effort to assess the contribution of 35 core courses to these three educational outcomes, faculty teaching core courses were asked to rate the emphasis placed on three kinds of educational outcomes: knowledge, critical thinking skills, and intellectual curiosity. Emphasis on *integrated fundamental knowledge* was relatively equal across all four academic divisions. Unit-weighting the emphasis on knowledge made divisional differences in relative emphasis on critical thinking and intellectual curiosity more apparent (Porter, 1997). The table below shows that faculty teaching Basic Science and Engineering core courses reported placing much less relative emphasis on student skills and attitudes. The assessed contributions shown in the right half of the table were drawn from a comprehensive study of the contributions of 35 courses in the Academy's core curriculum. Seven interdisciplinary faculty teams considered a wide range of standardized inputs from students, faculty and course syllabi to determine each course's contributions to students' mastery of integrated knowledge, ability to frame and resolve ill-defined problems, and intellectual curiosity (Porter, 1997). Contributions were rated on an absolute 7-point scale with 4 being neutral.

	Relative Emphasis by Academic Division (based on faculty self reports)			**Avg Rated Contribution to:** (7 pt scale - 7 asmnt teams)		
	Knldg	**Skills**	**Attds**	**Knldg**	**Skills**	**Attds**
Basic Sciences	1.0	54%	63%	4.88	4.66	4.68
Engineering	1.0	85%	69%	5.03	4.40	4.36
Social Sciences	1.0	94%	87%	5.36	5.56	5.42
Humanities	1.0	114%	108%	4.79	5.06	4.94

Since courses were assessed individually, the results could also be used to identify which pedagogical practices were associated with contributions to the three outcomes. The extent to which a course involved group work was negatively associated with the course's contribution to students' attainment of integrated knowledge and increase in intellectual curiosity. Closer examination showed that this was especially true in engineering core courses where the correlation between the proportion of group work and assessed contribution to students' knowledge approached -.80. The proportion of the student's course grade that depended on computation was negatively related to its critical thinking contribution (but showed a slightly positive relationship to knowledge attainment). Technical courses specially designed for *non-technical* cadets were generally found to contribute the least of all core courses to any of the three outcomes - these were also the courses that tended to employ student groups the most (Porter, 1997).

Other pedagogical process variables have also been found to relate to student learning. Although competition broadens performance distributions and allows selection among students to occur with greater confidence, its net effect on learning is generally negative (Kohn, 1986). Students also learn best when they are allowed to make mistakes, identify, and correct them. Overemphasis on external contingencies (either rewards or punishments) is likely to leech satisfaction and pride of ownership from learning (Glasser, 1990). Ironically, learning becomes much more likely when faculty simply "lighten up" and "let it happen". In classrooms with coercive climates, feedback about student learning is likely to be resisted and resented. In contrast, in classes with more collegial climates, feedback is likely to be used to enhance understanding, on both sides of the podium.

***Class content** - How can faculty best judge the utility of their class content choices? (How do we know if we are teaching the right stuff?)*

In my opinion, "*how to teach*" is a much more significant question than "*what to teach*". In fact, once the *how* is mastered, practically anything can be taught and learned. However, *what* to teach *does* matter, and careful consideration from many perspectives is appropriate. Material should be relevant - the more salient the connections to "real life", the more likely the material will be intrinsically motivating to students (and the less coercion will be required to get them to study). Course content should connect to what students already know as well as what they desire to learn. To the extent students perceive the course material as a bridge between their present situation and the attainment of their aspirations, little external pressure is required to motivate them to study. This is not a recommendation to "lower standards", "reduce rigor", or "pander to the lowest common denominator". Students take pride in accomplishing challenging tasks if they recognize the relevance of the material and know they'll receive the support needed to be successful. The particular content of courses should be determined by the current understanding within the discipline and its range of applications to business, industry, and education.

In order to appreciate the potential contributions of particular disciplines and perspectives, students need to learn the "stories" that structure the discipline. For this reason, teachers who are naturally inclined to consider "the big picture" rather than prematurely focus on details are likely to be seen as being more effective educators by their students (Porter and Benson, 1995). This preference is reflected by the i**N**tuitive vs. Sensing dimension of the Myers Briggs Type Indicator. The correlation between a faculty member's preference for intuition over sensation and rated effectiveness is about .30. In comparison, differences in teachers' Introversion or Extroversion preferences show nearly no association with students' assessment of teacher effectiveness or students' learning.

Pedagogical issues *- How can faculty assess the goodness of fit between student learning goals, pedagogical methods, and student learning styles? (What kinds of classroom feedback best guide changes in faculty teaching?)*

It is essential that institutional purpose, policy, and pedagogy align with one another (Porter & Light, 1994). For example, if the purpose is *development*, then collaborative approaches that encourage students to teach one another are appropriate. However, if the purpose is actually *selection* through differential performance, then competition is a more appropriate paradigm. Faculty themselves must have a coherent story; they should know what they want to accomplish and should have developed a range of activities that support the accomplishment of these objectives (Angelo & Cross, 1993). They must also regularly check progress toward these goals.

Education is not a ballistic process; it is much more like sequential hypothesis testing. Each lesson plan is a hypothesis - it is derived from the teacher's mental model of the way the world (or at least the classroom) works. During the execution of the plan, feedback from students should be collected and considered frequently. Classroom assessment techniques such as the 3 main (or "murkiest") points, quizzes, verbal responses, and student questions, all provide relevant information (Angelo & Cross, 1993). For most of us, our models need to be adjusted as evidence disconfirms some of our initial assumptions. If we truly endeavor to model learning, this should be a cause for celebration rather than chagrin.

It is important to recognize and appreciate that individuals differ from one another. There are likely to be at least several viable approaches to any particular problem, especially ones relevant to the "real world". A diversity of perspectives can enliven classroom discussions and enrich learning. Diversity creates a constant tension and provides the impetus for students and faculty to engage in collegial conversations with the goal of understanding each other's perspective.

However, the influence of individual learning style differences on most general educational outcomes is very small. It might even be argued that it is more important that students develop skills contrary to their natural affinities. Fortunately, faculty do not have to develop individual syllabi for each student. What they can do is create a syllabus rich with options and opportunities for students to master the material and acquire important individual and group skills in the process. The positive effects of providing choices and respecting a variety of student abilities are far greater than precisely matching academic tasks to student temperaments. Late adolescence is a volatile time; the reliability of most learning style instruments does not support reliance on categorizations determined early in the semester, let alone those acquired during freshman orientation.

Risks & costs of innovation *- What are the risks (apparent or real) of classroom innovation? What does it take for individuals to undertake them? How can assessment reduce them? What are the costs of classroom innovation? Who bears them? How can assessment address them?*

Every system is perfectly designed to yield the results observed. If perceived risks are sufficient to stifle innovation, the institution's days are numbered. Competition among educational institutions is increasing; there are more routes to a wider variety of degrees than ever; institutions that do not adapt will wither. *To innovate or not to innovate,* is that really the question? If an innovation succeeds, teachers, students, and the institution win. If an innovation fails, valuable information is gained and once again students, faculty, and the institution benefit. However, if innovation is absent, nothing is learned; alternatives disappear and everyone loses. Innovation creates the variability necessary to assess effects, consider alternatives, understand processes, and improve institutions.

However, innovation for its own sake can be dangerous: there is a difference between a *vision* and a *hallucination*. Effective assessment can help distinguish the two (Porter, 1997).

Bureaucratic traditions that developed in an environment of shrinking budgets and intramural competition encourage inappropriate and ineffective approaches to assessment. Certain *cosmetic* approaches to assessment attempt to insure positive appearances at the expense of gaining insight into processes and effectiveness. Such pseudo-assessments hide variance and obscure causality in their rush to proclaim perfection. The *sine qua non* of assessment is the same as for science: *we must endeavor to disprove our own assumptions*. In an educational community, where trust and understanding are valued and individuals are respected for their commitment to the common enterprise of learning, institutional and classroom assessment can contribute to the educational process. Interestingly, the keys to creating learning organizations are very similar to those for enhancing student learning: authenticity, understanding and trust.

Author's Note: Opinions presented in this paper are those of the author and do not necessarily reflect those of the U.S.A.F. Academy, Department of Defense, or any other government agency.

References

Angelo, T.A. & Cross, K.P. (1993). *Classroom assessment techniques: A handbook for college teachers.* San Francisco, CA: Jossey-Bass Publishers

Atkinson, R.C. & Raugh, M.R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 126-133.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Craik, F.I.M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology, General, 104*, 268-294.

Glasser, W. (1990). *The quality school: Managing students without coercion.* New York: Harper & Row.

Gould, S.J. (1981). *The mismeasure of man.* New York Penguin Books.

Kohn, A. (1986). How succeed without even vying. *Psychology Today.* (Sept) 22-28.

Kosslyn, S.M. (1981). The medium and the message in mental imagery. *Psych. Review, 88*, 46-66.

Palmer, P. (!998) The Courage to Teach. San Francisco: Jossey-Bass

Pavio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rhinehart and Winston.

Porter, D.B. (1991). A perspective on college learning. *Journal of College Reading and Learning, 24-1*, 1-15.

Porter D.B. (1997). *United States Air Force Academy Educational Outcomes Assessment Working Group, Phase 2: Final Report.* USAF Academy Technical Report 97-2.

Porter, D.B. & Benson, M.J. (1995). Correlates of course and faculty perceived effectiveness. In D. Porter (Ed.) *United States Air Force Academy Educational Outcomes Assessment Working Group, Phase 1: Final Report.* USAF Academy: Technical Report 96-4.

Porter, D.B. & Light, J. (1994). The institutional integrity of educational systems. In *9th Annual Conference on Quality and Assessment in Higher Education, Commissioned Papers.* Washington, D.C.: American Association of Higher Education.

Rogers, T.B., Kuiper, N.A. & Kirker, W.S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*, 677-688.