



EMBRY-RIDDLE
AERONAUTICAL UNIVERSITY

College of Career Education

Tenth Annual Symposium

on

Teaching Effectiveness

Evaluating Student Performance

Presented by the Faculty of Extended Campus

October 30, 2002



Table of Contents

	Page
Introduction.....	ii
Foundation of Academic Assessment By James A. Messerschmidt.....	1
Evaluating Student Performance in University Level Course Work By Melvin Anderson.....	7
Measuring Outcomes of Students' Learning Involving the Learning Modalities, Domains, Critical Thinking Skills Levels, and Right-and Lef-Brain Thinking By Frances Johnson.....	18
Peer Evaluations, and Team Learning in Undergraduate and Graduate Education By Bruce A. Rothwell.....	39
Writing As A Tool for Teaching and Evaluating Student Performance By Nina Haydel.....	55
Writing Is A Verb, Not A Noun By Patricia Valley.....	70
Legal Crib Sheet; Promoting Levels of Processing and Learning By Mary Jo Smith and S. Jay Samuelson.....	85
Cognitive and Affective Domain Learning Assessment Choices By Ronald Clark and Joseph Price.....	96



Introduction

Our topic for the Teaching Effectiveness Symposium for 2002 centers on evaluating student performance. The Faculty Development Committee was particularly interested in receiving papers related to assessing student performance in areas that fall within the affective domain.

Evaluating student performance serves several purposes. Obviously, it is a vehicle for assigning grades to students. More importantly, by obtaining a measure of student performance we are able to assess whether we are meeting our objectives, the objectives of the institution, and the objectives of the student. We are better able to judge our own effectiveness as teachers and we are better able to assure the public of our progress in educating those we address.

To the extent that we do not evaluate well, we negatively impact our ability to improve and strengthen our teaching. Evaluating student performance, then, is much more important than just being able to attach a fair and equitable grade to a student's name, or to be able to justify our actions upon an appeal of grades. It is to a large degree the marker that we use to judge or evaluate our performance.

The Faculty Development Committee of Embry Riddle Aeronautical University hopes that you will find these ideas from your colleagues both helpful and enjoyable.

These papers were selected by a jury of peers, in blind review.

The views expressed in the papers presented are those of the authors. These views do not necessarily reflect the views of the Embry Riddle Extended campus Faculty Development Committee, The Extended Campus administration or the University.

Jury:
Dr. Steve O'Brien
Dr. Dan Johnson
Dr. Bruce Rothwell
Dr. Ron Clark
Dr. Ron Fitzgerald

Editor and host: Dr. Earl Wheeler
Proceedings production: Ann Fagan
Facilities and arrangements: Lil Hickman



Foundation of Academic Assessment

James A. Messerschmidt

Embry Riddle Aeronautical University

Abstract

Student assessment is being impacted by the global economy concept. Students will be required to demonstrate inventive thinking and problem solving skills. This will require educators to rethink how to assess student's performance. Traditional assessment that focus on subject content and consists of objectively scored criteria referenced tests will not meet future assessment requirements. Performance assessments will adequately measure many of these higher level thinking skills. The foundation of performance assessment is the performance objective that must possess validity. This validity is determined through a thorough needs assessment that results in educational goals that define the performance objectives. These objectives will then be the source for performance based assessment.

Foundation of Academic Assessment

Student performance assessment is rapidly changing as educational institutions face challenges largely brought about by the new global economy concept. This emphasis on a global economy has established a requirement for students to move into the market place with more inventive thinking and problem solving skills than students of past decades (CEO Forum, 2001). In order to meet these new challenges, students will be required to not only understand the basic concepts and theories but also use higher order thinking and problem solving skills. Students will be required to be information literate. Doyle (1992) defined information literacy as "the ability to access, evaluate, and use information from a variety of sources." (p. 2) From an assessment standpoint, this requires educators to rethink how to assess to student's performance.

Definition of Assessment

In its general usage, assessment is the process of systematic gathering, interpretation, and use of information about student cognitive, behavioral and attitudinal outcomes for purposes of improvement. It determines a rate or amount and is used as an activity to measure student's change in behavior (learning) and other human characteristics. According to Smith and Ragan (1999) assessment serves two purposes: 1) provides information on student's performance and 2) how well the instruction is designed and/or presented.

Traditional assessment normally focuses on subject content and consists of objectively scored criteria referenced tests consisting of multiple-choice, true/false and/or short answers questions. This type of assessment focuses on the acquisition of and the memorization of factual information. It can be argued that this type of assessment cannot adequately measure many higher level thinking skills nor can it be used to measure many other important institutional goals (Baird, 1997). On the other hand, performance based or authentic assessment deemphasizes specific content, focuses on broader outcomes and is grounded in the authentic, real-life activities conducted in the classroom and/or laboratory (NcNeir, 1993; Armstrong, 1998). It requires students to construct a response, create a product and/or perform a demonstration (DODEA 2002).

Foundation of Assessment

The foundation of academic assessment is the performance objective. The performance objective defines the outcomes that students are required to perform at the completion of the instruction. Most traditional instructional development models are in agreement that instructors should develop their assessment tools directly from the performance objectives. This holds true regardless of the type of assessment or whether the instructor is measuring lower order thinking skills defined in Bloom's Taxonomy (Bloom, Mesia and Krathwohl, 1964) as observations and recall of information; or higher order thinking skills that

require the comparison and discrimination between ideas, assess value of theories and presentations, make choices based on reasoned argument and verify value of evidence.

If the performance objective does not reflect the level of behavior that the student is expected to perform at the completion of the instruction the assessment process will be flawed. In the case of performance assessment, it must possess *concurrent* or *predictive* validity. That is, how well the student will perform beyond the classroom and in the real world (Eisner, 1999). According to Wiggins (1994) assessment can only be an integral part of the learning process when it reflects the learning objectives. As an example, if the instructor develops a performance objective that requires the student to demonstrate a problem solving behavior then the student should be assessed on his/her ability to solve a problem. Often however, the student is not assessed on this behavior. He/She is given a criterion referenced examination that only measures lower order thinking skills.

Identifying the Need(s)

If the performance objective is the foundation for assessment, then it is extremely important to understand how it should be developed. The first step in developing performance objectives is to identify if there is a real need for the instruction. Dick and Carey defined need as "a gap between what is and what should be." (p. 15) Concurrent or predictive validity can only be established by conducting a thorough needs assessment.

Often performance objectives are developed in a vacuum without a clear understanding of the need(s). As a result, the gap that was first identified is never closed. Therefore several different parties should be involved in identifying the need(s). All curriculum and teaching decisions should be based on the need(s) and how best to facilitate the desired outcomes. These outcomes should be a consensus derived from the community or organizational vision of the skills and knowledge students are required to master in order to be effective adults and/or productive workers. It focuses the instruction on ensuring that students master those outcomes and it asserts that all students can succeed. In higher education it should be, at a minimum, the institutional educators and industry that will be hiring the graduates defining the need(s). If the educator works hand-in-hand with the industry and vice-a-versa, the need(s) can be accurately identified (McNeir, 1993).

Develop the Educational Goal

The next step is to identify the educational goal(s) that are based on this assessment. Educational goals are very broad, generalized statements focusing on what the student should be able to do when the education is completed. It is important to understand that the goals setting process may be influenced by more than just need(s). Instructional design often takes place in a specific context. This context could include political, economic, technical or academic issues. The selection of any educational goal

should be done in terms of the following concerns:

1. Are there sufficient people and time to develop the instruction for this goal?
2. Are the goals acceptable to those who must approve this instructional development effort?
3. Will the development of this instruction solve the problem that led to the need for it? (Dick and Carey 1990)

Goal Domain

There are generally three domains in which goals can be performed and it is extremely important that considerable thought be given to identify the domain in which the goal is. The domains can include the: intellectual, psychomotor, or affective domain. The intellectual domain requires the student to demonstrate a unique cognitive activity, such as, problem solving. The psychomotor domain includes both mental and physical activity. Maintaining an aircraft at a specific altitude and airspeed is a good example of behavior in the psychomotor domain. The most difficult domain to assess is the affective domain. This domain includes behavior that requires the student to make choices and decisions. The problem with assessing a behavior in this domain is that change does not occur in a short time. Therefore affective domain types of goals will be long-term goals that cannot be measured until the completion of all of the instruction. The

assessment of this type of behavior can include the measurement of both psychomotor and cognitive skills (Dick and Carey 1990).

Developing the Performance Objective

From these instructional goals performance objectives are developed. Performance objectives were popularized by Robert Mager in the 1960s. A performance objective indicates the observable behavior that a student will do to demonstrate that the lesson has been learned. The verb used must be an action verb that is measurable (observable). Performance was described as "an intended result of instruction rather than the process of instruction" (Mager, 1984, p. 5). The tradition performance objectives contain the following:

1. Performance: an objective always says what a learner is expected to be able to do.
2. Conditions: an objective always describes the important conditions under which the performance is to occur.
3. Criterion: wherever possible, an objective describes the criterion of acceptable performance by describing how well the learner must perform in order to be considered acceptable.

When using performance assessment, there are three key features that should be considered.

1. The student constructs the responses rather than selecting it.

2. The assessment format should allow the teacher to view the student's behavior that is based on real world requirements.
3. Establish scoring that reveals a pattern in students learning and thinking (Fuchs, 1995)

Connecting the Performance Objective to Assessment

It is quite easy to connect performance assessment to the performance objective. However, it is extremely important that the assessment is valid and measures the intended learning outcomes. This outcome depends upon the needs/task assessment that was completed early in the instructional design process. The task must be tied to real-work performance that reflects workplace and everyday activities. It is also critical that a realistic environment be identified in the conditions of the performance objective. This will help tie the behavior back to the original tasks that was determined from the needs assessment.

Defining the Aviation Industry Needs

The field of aviation education lends itself to performance assessment. Aviation universities must produce students that not only understand basic concepts and theories but they must possess higher order thinking and problem solving skills. Their success is driven by the aviation industry. As a result, aviation educators must continually define the needs of the aviation industry. An educator cannot

identify the need(s) in isolation nor can the industry expect the educator to know their need(s). This requires educators and aviation industry representatives to work together to define the need(s). Once the need(s) is/are identified, the educational institutional can then define their educational goals. These goals then will determine the students' performance requirements and only then can the instructor develop valid performance objectives. These objectives will then be the source for performance based assessment.

References

- Bloom, B. S., Mesia, B. B., & Krathwohl, K. R. (1964). *Taxonomy of educational objectives (two vols: The affective domain & the cognitive domain)*. New York. David McKay.
- Braid, H. (1997). *Performance assessment for science teachers*. Utah State Office of Education
CEO Forum on Education and Technology. (2001, March). *Education technology
must be included in comprehensive education legislation*. A policy paper by
the CEO Forum on Education and Technology.
- Dick, W. and Carey, L. (1990). *The Systematic Design of Instruction, 3rd ed.* New York:
Harper Collins Publishing.
- DODEA *glossary of terms* (March 12, 2002). Retrieved June 3, 2002 from DODEA
glossary of terms: www.odedodea.edu/instruction/curriculum/glossary.htm.
- Doyle, Christina S. (1992, June). *Outcome measures for information literacy within
the national educational goals of 1990*. Final report to the National Forum on
Information Literacy. Summary of findings. (ERIC Digest ED 351033).
- Eisner, E. (1999). *The uses and limits of performance assessment*. Phi Delta Kappan
Electronic Journal. Retrieved June 2, 2002, from www.pdkintl.org/kappan/keis9905.htm
- Fuchs, L. S. (1995). *Connecting performance assessment to instruction: A
comparison of behavioral assessment, mastery learning, curriculum-based
measurement, and performance assessment*. (ERIC Digest ED381984).
- Mager, R.F. (1984). *Preparing instructional objectives, revised 2nd ed.*
Belmont, CA: Lake Publishing.
- McNeir, Gwennis. (1993). *Outcome-based education*. (ERIC Digest ED363914).
- Smith, P.L. & Tillman J.R. (1999) *Instructional design, 2nd edition*. New York. John Wiley & Sons, Inc.

**Evaluating Student Performance in University Level Course Work:
The Certification of Academic Accomplishment Reveals
a Hidden Conflict in Academe**

Dr. Melvin Anderson

Abstract

Although higher education has begun to come to grips with problems in the process of evaluation of student performance, these problems continue to challenge teachers and universities. Grade inflation is rampant; teachers are seemingly unable and often discouraged in their efforts to return the grading system back to a former status when "C" was an average grade and "A" was reserved for the few truly exceptional students. Most teachers know that students expect high grades and will use a variety of methods to bring pressure on teachers who grade more objectively. To complicate the process further, students have access via the world-wide web to a wide variety of term paper and thesis web sites that make plagiarism very easy and very attractive—at the expense of legitimate research and learning.

This paper provides a framework for examining the effects of these problems and identifying a hidden conflict that causes them to persist. At the crossroad of a number of critical educational paths lies a decision point wherein teachers must decide whether to avoid conflict with students by awarding mostly high grades, or avoid the inevitable conflict with students' current and future employers and clients by awarding objective grades that certify actual learning and ability. As long as this is seen as an "either-or" relationship, teachers will opt out of conflict with students, or at best, compromise and satisfy neither alternative. The conflict is supported by the flawed assumption that there is no way to avoid conflict with the student except to award inflated grades.

The paper concludes that there is a third alternative that permits both requirements to exist—satisfying students realistically as well as satisfying their present and future employers—and in so doing, allowing the university to grade objectively and retain its academic reputation. This third alternative makes use of resources that are already available, and brings all three interests—students, employers and the university—together in a relationship involving professional societies and organizations to enhance genuine learning and evaluate student performance realistically.

Higher Education at the Crossroad

American higher education has come to a crossroad—a busy intersection—that brings a number of critical paths together at the same place and time. Approaching this junction is the explosion of high technology in the delivery of academic courses that includes the internet, video tapes and sophisticated classroom equipment. Interactive teaching methods are also coming down an avenue of development as teachers apply new technologies along with computerized teaching tools and simulations to their learning environments. Another path leading into this crossroad is the reasonable expectation in industry and the professional world that a college graduate whose credential—the college degree—represents real learning that can become real value to an employer or client.

The goal of higher education is preparing people for life. The objective of an effective process of evaluating student performance must support this goal by upholding the academic reputation of the institution. Therefore, a college degree must be considered a certification that the student has acquired skills, knowledge and understanding related to the discipline identified by that degree.

In many disciplines, even a college degree is insufficient certification. Additional certification and licensing are often required before a person can work or conduct business—another indication that student evaluation contributes not only to a certificate but also to actual certification.

At the same time, a sequence of ongoing problems has emerged within the nation's schools, some generated internally and others promoted by external influences that involve pre-college as well as post-college factors. These external influences seem to be creating the greatest internal problems for educators, and many of these problems involve the process of evaluating and reporting student performance. Despite increasing discussion and concern about improving the process of evaluating student performance in our colleges and

universities, efforts to implement effective student evaluation have become frustrated in recent years.

Pre-College Evaluation and Certification

Enrollment in a university is not an automatic administrative event. It requires legitimate secondary school qualifications—a high school diploma or successfully passing GED (General Education Development) Tests. One out of every seven high school diplomas issued each year is based on passing the GED Tests. Some states provide standardized testing before awarding a high school diploma. The New York State Board of Regents, for example, requires formal testing in English, Mathematics, Science and Social Studies for high school diploma certification.

Students entering college with valid high school diplomas are in effect certified by their states, schools or testing agencies to have achieved specific knowledge levels in their curricula. This, along with accreditation of schools, provides a good measure of assurance that a high school graduate is qualified for higher education.

Graduate Certification

The same certification concerns exist for students planning to enter post-graduate college degree programs. There exists a very rigid procedure at all universities for ensuring that students enrolling in master's degree programs have completed all the requirements for the bachelor's degree they claim, as well as specific prerequisites for advanced degrees they seek to earn. Most graduate schools require that prospective students complete standardized entrance exams given by private agencies such as the GRE or GMAT. Graduate schools also maintain standards regarding the accreditation of the colleges and degree programs of their prospective students.

Accreditation is not an automatic condition. The North Central Association of Schools and Colleges rescinded the accreditation of The University of Northern Colorado just over twenty years ago because of a number of substandard academic practices including

evaluation of student performance. The university expended considerable energy and money to correct these deficiencies and regain accreditation.

It is therefore evident that higher education has a fully developed process for ascertaining the certification of the credentials of students entering the university, as well as those advancing to higher levels within degree programs or enrolling in advanced degree programs. Certification is a valid and recognized process throughout higher education in America.

Evaluation is an Essential Part of All Certification

Certification is not just a word. From kindergarten through Ph.D., students are certified for promotion, advancement, diplomas or other certifications of educational accomplishment via diplomas, letter grades and standardized numerical values. Every accredited university reports academic achievement in a recognizable grading structure (usually A through F) with standardized credit hour measurements and numerical values (4 for an "A," 3 for a "B," etc.). Students' overall performance is calculated as an average of their earned grades, and is often held to specific minimum levels for remaining in good standing or being accepted for higher levels and advanced degree programs. And yet grading is currently one of the most controversial, disturbing and frustrating topics in virtually every level of education.

Grading Controversy Extends Outward

The current interest in grading and evaluation methods is ample proof that there is a great deal of ongoing concern about it. Authors and education experts generate literally thousands of books, articles and papers each year on evaluation of student performance. Educators are genuinely upset about the apparent widespread disparity between student performance and how that performance is measured and reported.

All university teachers have personal grading theories and methods, mostly in line with the policies of their institutions, and usually producing considerable dissatisfaction for teachers and students alike. We attend seminars and symposia; we discuss evaluation of student performance. We write papers and books on the subject and we develop personal or organizational formulae for doing it. We agree that a formal evaluation process is a necessary condition—a teaching tool for inspiring (or coercing) students to perform better in their coursework. We agree on the purposes and general content of the evaluation system and the kinds of problem we encounter with it, but most of us are very hesitant to try to change anything.

Standards and Measurements

Whatever evaluation concept a teacher uses is based on some set of standards and measurements derived objectively from students' attendance, participation, exam scores, presentations, research papers and the like. A good grading schema must begin with clear definition of what is expected and how it will be measured. Virtually every university course is built on a published Course Description, accompanied by specified course goals, textbooks and learning outcomes that teachers are required to address in their course content. Achieving these objectives requires a clear set of measurable performance outcomes in syllabi and student expectations. In all cases, the evaluation process is something internal to the system, and whatever evaluation methods are applied are supposed to be as prescribed by the policies of each institution. And although every teacher has a personal grading philosophy that s/he believes is appropriate, it is very unlikely that any single approach provides a perfect formula for fair and objective evaluation. But there is enough common ground on which most of us agree.

It is even more unlikely that any one study or paper will provide a formula that satisfies a majority of the academic community, yet we agree on the general aspects of what indicates how much a student has learned.

Whether we evaluate input factors such as attendance and participation, or output factors such as skills and understanding, there is a general confluence of viewpoints about how these should be measured and graded.

Trouble on the Report Cards

Despite general agreement in academe about the importance of evaluation and the apparent determination of most teachers to do it right, the grading process has become severely inflated. "A" has become the median grade. For this reason, more than all others, we are now searching for solutions that will re-create a central tendency in grading averages. If it were up to the teachers en masse, evaluation of student performance would suddenly return to the days when "C" was average, "B" was considered a good grade and "A" identified truly outstanding work. We know how to do it, but we don't.

Harvard University professor Harvey Mansfield is a long-time critic of grade inflation who offers some insight into the internal aspects of this phenomenon. In an article in the April 6, 2001 issue of the *Chronicle of Higher Education* entitled "Grade Inflation: It's Time to Face the Facts," he offers some thoughts that bring out the seriousness of the problem:

In a healthy university, it would not be necessary to say what is wrong with grade inflation. But once the evil becomes routine, people can no longer see it for what it is. Even though educators should instinctively understand why grade inflation is a problem, one has to be explicit about it.

Grade inflation compresses all grades at the top, making it difficult to discriminate the best from the very good, the good from the mediocre . . . some of my colleagues say that all you have to do to interpret inflated grades is to recalibrate them in your mind so that a B+ equals a C, and

so forth. But the compression at the top of the scale does not permit the gradation that you need to rate students accurately . . . mere recalibration does not address the real problem: the raising of grades way beyond what students deserve . . . we have lost the notion of an average student.

Mansfield connects the grade inflation problem with external problems like student course evaluations, student expectations rather than their own criteria, and the resultant effect on authority and morale:

The loss of the notion of average shows that the professors today do not begin with their own criteria for the performance of students in their courses. Professors do not say to themselves, "This is what I can require; anything above that enters into excellence." No. With an eye to student course evaluations and confounded by the realization that they have somehow lost authority, professors begin from what they think students expect. American colleges used to set their own expectations. Now, increasingly, they react to student expectations.

Thus another evil of grade inflation is the loss of faculty morale that it reveals. It signifies that professors care less about teaching. Anyone who cares a lot about something is very critical in making judgments about it. Far from the opposite of caring, being critical is the very consequence of caring. It is difficult for students to work hard, or the professor to get them to work hard, when they know that their chances of getting an A or A- are 50-50. Students today are still motivated to get good grades, but if they do not wish to work hard toward that end, they can always maneuver and bargain.

Mansfield concludes by admitting that the reasons for growing grade inflation are not entirely clear, but that the solution will come only when we put our standards first and muster the determination to act.

We must remember that learning really occurs in our institutions and we are capable of evaluating student performance objectively. And so why is the grading system inflated? Since it is not an internally-caused problem, then it must be an externally-caused problem. In other words, despite our own abilities and intentions, something is pulling our red pencils upward as we measure student performance and report it. It cannot be pre-college external factors such as high school certifications, because these are in the book long before students receive their first college semester grades. Therefore, it must be something external—something other than our own methods—that induces teachers as individuals to distort the evaluation process after we have set standards and transmitted expectations to students. Moreover, it must involve a hidden conflict that prevents or discourages objectivity despite our clear objective of implementing an effective student performance evaluation system and upholding a good academic reputation for our universities.

Two External Requirements

If we know that the stated objective is effective evaluation of students that upholds our academic reputation, then we also should know what requirements must be maintained in order to reach that objective. The two basic requirements are fairly obvious:

First, avoid conflict with students.

Students consider themselves customers who deserve something for their investment in tuition and effort: learning or good grades, or both. Whether students really seek to achieve academic objectives or are more interested in high grades is important because it defines what they consider proper in the evaluation process. Within the MTV generation we will continue to find both kinds of student.

Second, avoid conflict with the students' employers and clients.

Students' employers and clients, present and future, expect something for their investment if they pay their employees' tuition or trust them to provide specific skills their degrees indicate.

Both of these requirements must be met in order to achieve the stated objective. If we conflict with our students, we certainly will not uphold our reputation with them and we will probably lose many of them to "looser" schools or programs. If we conflict with the students' eventual employers and clients, we will probably lose their confidence in the institution's credibility and the credibility of its degrees and certifications.

In our litigious society, the university bears some of the responsibility for representing the capabilities of its graduates. It may even find it is liable for the actions of the people it certifies.

Avoiding the Obvious Conflict

In order to avoid conflict with those who place grades above learning, we tend to allow easy grading—pass almost everyone. Whereas students once earned grades, they often try to negotiate their grades. In some cases, they try to litigate grades by putting pressure directly on the teacher through formal complaint channels, persuasive techniques and a variety of outright coercive maneuvers. In rare cases, students take legal action against schools and teachers over grades. In that environment, teachers are only too willing to avoid conflict by easy grading. For many students, the grade is everything; for many teachers not interested in conflict, the grade is not worth fighting about.

Fanning the Fire

Evaluation is not a one-way street. Most schools obtain feedback about teachers and courses through a formal course evaluation process, usually one in which students complete a standardized form during the last session of a course. Students complete the forms anonymously, without teachers' inputs

or influence, responding to questions about various course "quality" issues. The forms typically contain objective multiple-choice questions, plus a number of open-ended questions requiring written responses. Completed forms are collected and delivered via administrative channels to data centers where numerical scores are computed. This is properly done before teachers award final course grades. Results, including the written responses, go to teachers' department chairs and perhaps deans, and eventually back to the teachers. Teachers know these evaluations are inevitable and they are concerned about ratings.

Teachers quickly learn what it takes to get high scores in these evaluations; happy students (happy customers?) write good evaluations. In this era of "credentialism," teachers cannot afford to overlook the importance of good formal evaluations *from* students, as well as good evaluations *of* students. Looking good is career enhancing.

Credentialism has become so dominant in our society that in some cases, people sometimes take the risk of claiming to have earned college degrees they really haven't completed. Recently, the United States Olympics Committee President Sandy Baldwin was forced to resign when it was learned that her resume contained false information about earned degrees. Baldwin was a talented and highly effective leader who fudged on her resume—and is now history. Her exit was appropriate, but we ought to consider how many others, who actually receive their diplomas, do not have the learning and talent needed to perform the jobs they were offered on the basis of their credentials.

The Quality Era

This concept of "quality" in higher education became popular during the onset of the "quality movement" around 1980. Manufacturers developed programs to measure and improve customer satisfaction, and universities began developing programs to measure and improve student

satisfaction. The results of these "quality" initiatives are now described in the history books of business and academe.

In the manufacturing arena, it eventually became evident that quality is important, but not the true goal of either business or higher education. Of the 30 American firms who have won the coveted Baldrige Award for Quality since 1987, six have already gone through bankruptcy—*after* receiving the award! Whereas the Deming Award for Quality in Japan focuses on product quality ("happy customers"), the Baldrige award is based heavily on input factors like innovative practices in human resource management—"happy employees," according to the Human Resources Learning Center. In academe, much of the quality we espouse is measured at the input level, before students have had a chance to apply what they've learned in the real world. At that level, all we have to go on is "happy students." It is becoming increasingly apparent that the "customers" of higher education are not just the students.

On the web

Another challenge has exploded into the college environment: plagiarism. In just two years, from 2000 to 2002, the number of "term paper" web sites has grown from 12,000 to more than one million. In June, 2002, the world-wide web could link a student to 1,094,805 sites from which students can buy ready-made term papers on any topic. All they must do is click on the keyword "term papers." These sites don't exist for aesthetic purposes; there is a market for such products in academe.

Plagiarism is not just a way for some students to get higher grades; it does irreparable damage to the students who don't do it, and in time, to the institutions that don't stop it.

Julie J. C. H. Ryan of George Washington University provides an incisive view of the issue of plagiarism and evaluation of student performance:

Often lost in the discussion of plagiarism is the interest of the students who don't cheat. They do

legitimate research and write their own papers. They work harder (and learn more) than the plagiarists, yet their grades may suffer when their papers are judged and graded against papers that are superior but stolen material. Students have a right to expect fairness in the classroom. When teachers turn a blind eye to plagiarism, it undermines that right and denigrates the grades, degrees and even institutions."

At the same time, a large number of plagiarism-fighting tools are available to teachers to help identify and counter ill-gotten term papers and reports. There are currently more than 194,000 web sites that offer software products and services for detecting and countering plagiarism. This cyber-jousting verifies further our awareness of the fact that credentials, not substance, are more important to many students.

Avoiding the Other Side of the Conflict

There are always two sides to conflict. The other side of the evaluation conflict deals directly with the certification that a college degree represents. To award a degree is to certify that a student meets the defined requirement for that degree. For example, a person with a degree in Aeronautical Engineering is expected to know something about airplanes, mathematics, physics, propulsion and flight dynamics. A person with an MBA is expected to know something about business management, finance, accounting, marketing, production operations, project management, information systems and the like. This kind of expectation rests in the minds of employers and clients who believe that college graduates are worth their salaries or fees because they have credentials—degrees—specifying the disciplines in which they have been educated.

Anyone who has waited in a physician's office has examined the medical degrees and certifications on the wall; we understandably want to know where the doctor learned his trade. We trust the

credentials of airline pilots because we know they must pass stringent requirements in flying schools and FAA aircrew evaluations. The same applies to other disciplines, even to music and the arts; a person either learns or doesn't learn to play the cello. Today, symphony orchestras desperately seek good cellists, while a million or more young people have guitars and dream of becoming rock stars.

Labor unions that once certified their members have opted not to be involved in evaluation activities. In recent years, union membership has declined significantly; their focus has now shifted to trying to preserve membership numbers. As a result, employers and clients of the various trades must look to other indications of competence, such as training and education credentials.

Many business and government organizations pay the college tuition for their employees, in the expectation that the expense will be repaid in the form of better and more valuable performance. If this performance improvement is not realized, these organizations may discontinue tuition assistance or avoid hiring graduates from specific universities. This is the conflict every university must avoid.

In order to avoid conflict with the present and future employers and clients of students, the university must legitimately discriminate between students who really demonstrate learning and those who do not. That means that as a prerequisite to avoiding conflict with employers and clients, grades must not be handed out to those who do not earn them. In other words, schools must maintain rigid grading and evaluation standards and be selective.

But maintaining rigid evaluation standards is in direct conflict with easy grading that avoids conflict with students. This is the hidden conflict that prevents effective evaluation of student performance as a process.

Breaking the Hidden Conflict

We must either be easy evaluators or objective evaluators. We cannot be both. We can either accept conflict with students by grading objectively, or accept conflict with their employers and clients by handing out unearned high grades. Or so it seems.

This conflict is held in place by a single flawed assumption: that there is no way to avoid conflict with the students except to award inflated grades. Because of this flawed assumption, we arrive at having to ask a question that has two wrong answers.

Are we more at risk of losing enrollments or of devaluing the image and marketability of our degrees? If one or the other must go, then the next task is to decide which it is. Or can we accept some kind of compromise wherein we raise the requirements a little and try to assuage the concerns of both sides? Experience tells us that compromise is usually a bad choice, except in politics. And so we eventually must come up with a third alternative. Otherwise, we continue down the more demanding path.

The Third Alternative

Teachers logically consider declining enrollments as a danger to their jobs and incomes. Moreover, teachers receive pressure to keep students in college from administrators and external factions that include teacher unions and other special interests with political influence. Every reduction in university enrollment is seen as a threat to someone.

In most cases, business concepts work well in higher education. The marketing model applies in the general sense, and colleges expend a lot of effort and money attracting new students to replace those who graduate each year. Unlike tobacco smoking, education is not addictive; nearly everyone quits before they die. Logically, every college web site and catalog focuses heavily on the reasons prospective students should choose this school or the other. With this as a primary motive, schools logically attempt to maximize enrollments and retain students. But motive alone cannot produce

the desired effect of an educational institution with a strong academic reputation. It takes effective methods and the means to deliver educational products to the end users (customers).

Who are the True Customers?

As long as the students are seen as the only customers of higher education, all the motives, methods and means will continue to focus on delivering that which the students dictate by their feedback in negotiating curricula, grades and faculty evaluations. In viewing the system from within, we tend to respond to immediate pressures from the students we see daily and consider customers deserving of being satisfied like people who buy hamburgers without regard for the fat content and its long-term health effects!

We are overlooking another aspect of these customers—the professional capabilities of the students we certify to their employers and clients. Some students' employers contribute to their education expenses, which means we have a quasi-legal responsibility to these employers and clients because we issue certificates of knowledge called diplomas and degrees. These are indeed our customers, although they tend not to make as much noise about what satisfies them.

In truth, the customers of higher education are twofold: the students who should benefit from their personal study efforts, and also those students' current or future employers and clients who pay for and expect to benefit from employing credentialed college graduates. Interestingly, serving one is in the long run serving the other as well. And yet they seem to be in conflict, and that to serve one is to abandon the other.

There is a third benefactor in this customer relationship: the university itself, whose long-term reputation (and perhaps its survival) inevitably benefits from an evaluation system that delivers what both sides of the conflict deserve.

Therefore, the only solution that can have long-term success is one that satisfies both sides of the problem: to avoid conflict with students and at the same time avoid creating conflict with the students' employers and customers.

The PTA?

Since 1897, primary and secondary schools have maintained a formal process in which students, parents, teachers and other primary and secondary school officials could interact to discuss specific educational issues and track students' general progress and performance in individual courses. Virtually every public and private school in America has an operating Parent-Teacher Association (PTA) or similar organization. Perhaps too many children today aren't benefiting from the advantages of the PTA because their parents don't participate, but the framework is in place for those who use it.

Unfortunately, when their offspring go off to college following high school, proud parents tend to breathe a sigh of relief because they don't have to go to any more boy/girl scout and PTA meetings! Someone else will take over the role of seeing that their little darlings go to school on time, do their homework and study for exams. Someone else will monitor what courses and programs they enroll in, and someone else will sign their report cards and admonish them when their grades aren't perfect.

Nothing could be farther from the truth! Nobody, except perhaps students themselves, will assume "parental" authority when students reach adulthood and go to college.

Rescue Points

Fortunately, there are some "rescue points" in the existing system that may save some students from academic disaster, if students will avail themselves of them. Rescue points are either places or events that enable students to clarify goals, or set new goals, or obtain encouragement and influence that will facilitate their academic growth. For example, there are a number of

clubs and professional fraternities, like Delta Sigma Pi for business students, which attract and encourage students with legitimate learning goals. Every university monitors its students' extracurricular activities in some way. Faculty and administrators still take time to meet with students and serve as advisors for their curricula and course selections, if students will take the time to see them.

Most of these rescue points are aimed at the students' and schools' immediate best interests. Very little is directed toward the other side of the conflict, which involves the students' current and future employers and clients.

If the PTA isn't around, and the rescue points don't address the potential conflict with students' employers and clients, then what action can be taken that formally links both student and employer concerns? It must be some "injection" or third alternative that does both.

The Third Alternative

The concept of the PTA is a good one, albeit focused on student-parent-school interests at the primary and secondary levels. But there is another interest that can replace the parents when students attend college, perhaps not as authoritarian as parents, but equally concerned about the students' learning outcomes and interested in the performance evaluation process that certifies students with degrees that imply specific abilities and professional skills.

Many students' current employers are paying tuition expenses as part of employee benefit packages. This is true of military, other government agencies and many private firms, large and small. These payers not only have a vested interest in how their education money is being spent; they also expect their people to return that investment in the form of more valuable performance. These students usually know their responsibilities and know that it takes more than a good grade to convince the employer that they have learned something in college. A "PTA-like" relationship is easily

Evaluating Student Performance In University Level Course Work

established in these cases, wherein the employers are directly involved in their employees' curriculum, course selection and academic progress, even though there are legally enforceable privacy concerns.

But what about the students who haven't yet entered the professional world? Many undergraduate college students have never held jobs, but are enrolled in degree programs aimed at specific professions. Most formal professions have professional societies that link people of the same profession through membership, seminars, annual symposia, local chapters and regular activities. A partial list of these national organizations includes the American Society for Quality (ASQ) with 70,000 members, the Society of American Military Engineers (SAME), the Society of Automotive Engineers (SAE) and the American Institute of Electrical Engineers (AIEE). We all know of the American Medical Association (AMA), the American Bar Association (ABA), the American Dental Association (ADA), the American Psychiatric Association (APA), the American Institute of Certified Public Accountants (AICPA), and the influence these organizations have had in academe as well as in the professional world. Virtually every discipline presented in a degree program in college can relate to a national organization that has nearby local chapters and members. And these chapters are already interested in what is being taught in colleges around the country that affect their future memberships.

Many professional societies have student chapters that look for future professional members at colleges and universities. And they are welcome in the colleges because they encourage enrollments and inspire better academic performance. Their presence, however, is not sufficiently formal and their influence does not often get into the process of evaluating student performance.

Making a Difference

"Credentialism" is the belief that one's paper persona is more substantial than the person; the belief that an inflated resume or degree

representing inflated grades is as good as state-of-the-art knowledge, experience, and ability to perform specific skills. Credentials may get some of us in the door, but they won't keep us on the job once the employers and clients discover that we can't do the work.

Professional organizations can help avoid the pitfalls of credentialism by having a direct influence on students' learning. They can provide regular feedback to educators about those they've graduated with flying colors who can't even find the flagpole. When this relationship is established formally, the evaluation of student performance becomes a legitimate reflection of actual learning and ability; grade inflation may then become an embarrassment to easy-grading teachers. Likewise, teachers will begin to use the tools available to detect and discourage plagiarism. Many students will begin to realize that both sides of the evaluation conflict are collaborating in the students' best interest by standing fast on what constitutes learning and objective evaluation.

With this re-orientation of student attitudes must also come a re-orientation of teacher attitudes. This will require, among other things, a redirection of the process of student evaluation of teacher performance. At present, this process is mostly a measure of customer satisfaction with students seen as the only customers. As long as teachers fear student feedback that criticizes teachers who require hard work, this evaluation tool will continue to work against achieving better student performance.

To activate this already-available "PTA-like" injection to a level that influences both teachers and students, we must make it an integral part of the planning, programming and evaluation processes of the university.

It will require leadership and the willingness to change.

Bibliography

Books:

Perelman, Lewis. *School's Out*. New York: William Morrow, 1992. The author offers some alarming predictions about the survivability of traditional institutions of higher learning in the new "electronic classroom" era.

Rackham, Neil. *SPIN Selling*. New York: McGraw-Hill Book Company, 1988. Presents a systematic series of questions designed to cause clients (students, in our application) to reveal their true intentions and provide teachers with the information needed to offer solutions that truly benefit students and the university.

Articles and papers:

Harvey C. Mansfield, "Grade Inflation: It's Time to Face the Facts." *The Chronicle of Higher Education*, April 6, 2001. This Harvard University professor offers some insight into the internal aspects of this phenomenon and some thoughts that bring out the seriousness of the problem.

Julia J.C.H. Ryan, *Student Plagiarism in an Online World*. Prism Online, December, 2001. A George Washington University teacher describes the enormous growth of online cheating by college students and some ideas on how to catch a plagiarist.

Alison Schneider, "Insubordination and Intimidation Signal the End of Decorum in Many Classrooms." *Chronicle of Higher Education*, March 27, 1998. A review of the causes and effects of growing problems in evaluating student performance.

James W. Michaels and Dirk Smillie, "Webucation--Some smart investors are betting big bucks that Peter Drucker is right about the brilliant future of online adult education." *Forbes Magazine*, May 15, 2000. A challenging discussion of Drucker's analysis of higher education, the "e-learning" race, and the differences between the various learning environments.

Robert Lenzner and Stephen S. Johnson, "Seeing Things as They Really Are--College Education as we Know it is Threatened (Peter Drucker)." *Forbes Magazine*, March 10, 1997. Another strong warning that higher education faces formidable challenges to its very survival.

Michael Prowse, "Endangered Species. Modern electronic technology could mean that the days of academics at higher-education institutions are numbered." *Financial Times*, November 20, 1995. A hard-hitting discussion about changing academic environments.

David Rothenberg, "How the Web Destroys the Quality of Students' Research Papers." *The Chronicle of Higher Education*, August 15, 1997. A sobering view of the misuse of today's electronic learning tools.

**Measuring Outcomes of Students' Learning
Involving the Learning Modalities, Domains,
Critical Thinking Skills Levels, and
Right- and Left-Brain Thinking
Frances Johnson**

Abstract

There is no paucity of literature that relates to the learning domains, the modalities, critical thinking skills, or right- and left-brain thinking. Charts for various aspects of the preceding aspects have taken the forms of pyramids, circles, and the more conventional square or rectangle as well as bubbles and diagrams. None, however, have attempted to chart sample evaluative vocabulary that links the critical thinking skill levels to both the learning domains and the learning modalities. Textual comments indicate links between the preceding and right- and left brain thinking. The extrapolation of the sample evaluative vocabulary provides a quick reference to provide effective and valid evaluation of student learning as follows:

1. Enhance the inclusion of test items related to right- and left-brain thinking;
2. Create a stronger link between test items and learning modalities;
3. Develop a stronger link between test items and the affective domain;
4. Improve the link between test items and the cognitive and psychomotor domains;
and,
5. Forge a stronger link between test items in the preceding categories and the Johnson Critical Thinking Skills Levels.

Measuring Outcomes of Students' Learning Involving the Learning Modalities, Domains, Critical Thinking Skills Levels, and Right- and Left-Brain Thinking

Read, review, and regurgitate! For both faculty and students, these words may generate the thought that this is the prevailing paradigm that governs the examination of skills presented in institutions of higher learning. That is not to say that for entry level courses one should avoid the delivery of core course content as well as the measurement of the students' acquisition of basic vocabulary and concepts before embarking on more ambitious academic journeys. However, choosing the appropriate method or methods for assessment or evaluation necessarily involves the consideration of a number of seemingly disparate factors. Such factors as short and long term memory, the effects on learning based on the dominance of either the right brain or the left brain, and the mental processes that have been identified as the cognitive, affective, and psychomotor domains come into play. In addition, the auditory, visual, and kinesthetic or tactile learning modalities have been stirred into the mix. The preceding as well as many variations of the preceding have been the targets of much speculation and research in the ongoing effort to enhance both the delivery of course content and its assessment or evaluation. Because the basic 3R assessment tool fails to address in a meaningful way either the psychomotor or the affective domains of student performance, tests, examinations and other assessment tools are subject to ongoing revisions and updates in order to provide a broader based and more accurate picture of

student progress. The preceding paradigm, although modest in its pretensions, serves reasonably well at the basic knowledge level to ensure that core knowledge has been mastered. The simplicity of the approach is that it has application for large entry level lecture classes that mandate that tests be scored by machine or with a template. However, it provides limited utility for assessment of skills at the higher levels of cognitive learning. Such a paradigm provides even less utility for assessing skills in the psychomotor and affective domains. As is so with the cognitive domain, the issue becomes that of identifying relevant concepts and of ranking them according to importance as well as the desired level of performance in order to determine an appropriate method of assessment. These constructs and domains permit discussion as well as some small measure of understanding of the manner in which the human brain functions. They also permit speculation about what is learned and about what is retained. Retention of what is learned is one goal of education. Utilization of what is learned and retained is another. As instruction progresses through Bloom's taxonomy and/or Johnson's (2001) Critical Thinking Skills Levels, assessment of student learning becomes more complex. For most college level courses verifying that basic knowledge has been acquired and retained does not provide a complete measure of a student's progress nor does it fully address a student's application of acquired knowledge.

If, as stated previously, verification (assessment) of knowledge acquisition, retention and utilization are important, then it follows that

the assessment of what is learned provides justification for what is taught. Ideally, assessment of students' learning can and should provide insights into the modification of methods for delivering course content. By providing a variety of techniques the instructor ensures that each student has the opportunity to exercise the domains and/or modalities with which he/she is most comfortable as well as to develop skills in the less favored areas. For a multitude of reasons, student progress must be examined with the same degree of professionalism that is utilized to deliver course content.

Retention of Learning and Learning Modalities

According to data published in several online sites that are maintained by colleges, universities, and the Tehama County (California) Department of Education, short-term memory is not very efficient under certain conditions. According to the learning pyramid attributed to the National Training Laboratories, Bethel, Maine, which depicts the methodology used and average learning retention rates, the learner who is relying on hearing or reading as the single method of processing information, the retention of learning rate is 5% and 10% respectively. When hearing and seeing (visual and auditory components) are used together retention of learning jumps to 20%. Demonstration increases the retention rate by about 10% bringing the rate to 30%. Participating in discussion groups brings the retention rate to 50%, practice by doing to 75%, and teaching others to 90%. Experience, simulated experience, challenging activities, and teaching others provide high rates of improvement in retention

when measured one month later according to data provided under the auspices of the New York Institute of Technology online sources. It is not completely clear who performed the study that provides the basis for the various educational entities that use the data although Ann Boultinghouse of Keystone Instructional Services attributes it to William Glasser. Nor was the date of the original study available. However, it is reminiscent of the SEE, SAY, and DO (visual, auditory, and kinesthetic) paradigm often used by those engaged in elementary education. In a scholarship program operating under the auspices of the New York Institute of Technology, it is indicated that retention one month later showed losses of learning and/or distortion when learning was dependent on just one of following: reading, hearing or seeing. Participating in challenging activities and teaching others provided the greatest amount of retention of learning one month later, 83% and 91% respectively. There is enough evidence to warrant consideration of the implications of retention of learning rates under varying conditions when choosing a method for the assessment of student learning.

Fox Valley Technical College emphasizes a pragmatic approach to curriculum assessment to the extent that it provides a Modality Study Aids web page with specific study techniques suggested for visual learners, for auditory learners, and for kinesthetic or "hands on" learners. Inherent in this is the implication that classroom assessment should include methods of evaluation that engage each of the learning modalities.

Right Brain, Left Brain, and the Affective Domain

Much of the course content that falls within the cognitive domain is oriented toward left brain processing at both the instructional and assessment levels. Both instructional and assessment tools often utilize Bloom's taxonomy. Through the inclusion of essential definitions, facts, sequence, concepts, symbols and other knowledge based criteria, fundamental or core knowledge is delivered to enhance the foundation for higher level critical thinking skills in the cognitive domain. Evaluation of such knowledge is necessary to ascertain whether students are prepared to apply the skills acquired. Hence, many examinations such as multiple-choice, short answer, short essay and essay answers are the assessment instruments of choice for measuring knowledge acquired. Words such as *define, label or name, describe, state, identify, choose and find* to list a few may appear in short answer and short essay examination questions. The preceding assessment tools fall into the time-honored tradition that provide results that, if well constructed, are viewed as reliable and valid.

As noted in the preceding paragraph such measurement tools test predominately left brain processes that have been labeled as *logical, sequential, rational, objective*, and occupied with *examining the parts*. Right brain thinking is often *random, holistic, intuitive, subjective*, and is focused more on the big picture rather than on details. Bernice McCarthy has extended the concept to include teaching to "the

four styles (concrete experiences, reflective observation, abstract conceptualization, and active experimentation) using both right and left-brain processing techniques. Ideally, assessment of a student's achievement has provisions for evaluating learning in both the right- and left-brain.

It likely is no surprise that right-brain thinkers often engage in activities that align closely with the affective domain. Nor is it likely to be a surprise that Vygotsky's theory of social development, Bandura's theory of self-efficacy and Maslow's theory of self-actualization also relate to the affective domain. The performing arts and creative projects are frequently viewed as primarily engaging right brain processing because of the impact on the emotions of the audience as well as for the affective behaviors inherent in performing such activities. Attitudes, values, beliefs, and emotions generally are activated as subjective or right brain thoughts. Betty Edwards (1989) in *Drawing on the Right Side of the Brain* states that intuition, imagination, and creativity are components of right-brain thinking. Measurement or assessment of the results of learning may include performance as well as attitudinal and/or value scales based upon a previously determined definition of affective behaviors and the circumstances under which they are expected to occur. However, if self-administered attitudinal and/or value scales are used, the results may be skewed by the disparity between what the subject believes to be happening and what others perceive as actually happening. Longitudinal studies and evaluations tend to provide a slightly more accurate

assessment of the development and retention of emotions, attitudes, and values within the affective domain. There are, however, obstacles to be overcome when attempting such assessments. Obstacles that can be anticipated include such things as the mobility of the society in which we live and the frequency of that mobility as well as whether people will respond and whether the response will be handled in a timely manner to such queries. Furthermore, there exist within the population those who with sincerity question not only whether we can but also whether we should attempt to influence the attitudes, beliefs, and values of students.

Yet, the basis of higher education is to promote, encourage, nurture, and teach attitudes, beliefs and values that will provide the foundation for a rewarding career and a rewarding life. Furthermore, your presence in this room is indicative that each of you at some point in time made a conscious decision based on your attitudes, beliefs, and values to expand your educational horizons. Your presence here is also an indicator that since that time you have actively participated in academic life by committing time and energy to research, instruction, assessing student learning and to program development and modification. Much of this has been done while participating in family and community life. In all probability, you have had or still have a direct tie to the aerospace industry.

Who motivated you to embark on an educational journey at a time in your life when you made that decision? Can you still name the individuals most responsible for your decision to

attend an institution of higher learning as well by what means the influence occurred? Or was your decision initially influenced more by economic concerns than for other reasons? The attitudes, values and beliefs involved in such decision-making fall within the affective domain although Huitt (1997) would classify the process that may have motivated you to engage in that behavior as *inculcation*. However, consider whether those who influenced you made a conscious decision to have you consider your options. If your answer is in the affirmative, it is a small step to accept the charge to provide opportunities for students to modify and/or develop attitudes, values, and beliefs which will sustain them in their many and varied endeavors.

Is it necessary to establish more course offerings devoted exclusively to the teaching of values and ethics? Or do courses in the humanities, psychology, philosophy, and religions provide opportunities to provide that which is needed? Should more resources be shifted to appreciation of the fine arts associated with the affective domain as well as to the strategies and techniques used by their creators? Or are some, if not all, of the preceding courses providing opportunities to reinforce and develop appropriate attitudes, beliefs, and values necessary for success in the workplace, in the community, and in private life? The preceding questions require the identification of what is being done in regard to the affective domain with the focus on the development of appropriate methods for measuring the outcomes of those efforts.

In addition to the preceding, it is beneficial to examine how the components fit into a paradigm that provides insight and focus with regard to assessing student progress in various areas utilizing a variety of modes of delivery of course content in a variety of settings. Course content is delivered in labs and classrooms, in the field, through distance learning, and through electronic media such as television and the Internet. Ideally, such a paradigm should address recognized domains, critical thinking skills, learning modalities, and right and left brain functions. Desired levels of achievement and retention of learning are governed in part by course outlines and mandated acceptable levels of proficiency deemed necessary to advance to the next academic level. However, carefully planned and well written classroom objectives become a complement to course outlines and course objectives.

Posing fewer problems is the assessment of the psychomotor domain which, in an ideal situation, is performance based as is likely to occur with the performing arts or when presenting speeches in speech courses or demonstrating skills in physical education classes or conducting experiments, completing labs, etc. Although the tactile or kinesthetic modality is the primary modality for the psychomotor domain, the visual and auditory modalities as well as the cognitive (left brain) and affective (right brain) domains are involved in the learning process. It is suggested that the preceding considerations need to be included in the assessment of student learning as well. In fact, the cognitive aspects that are engaged in psychomotor development are tested

more frequently than performance itself.

However, the most noticeable exception to any attempt at meaningful assessment of student learning in the psychomotor domain is on normed, standardized tests. Performance testing on such a scale is cost prohibitive according to figures released during the spring of 1996 by the General Accounting Office. The estimate for a "national multiple-choice achievement test" is approximately \$42 million, "while a slightly longer test with short, performance-based questions would cost \$209 million."

Finally, given the time constraints imposed by the length of sessions used by the Embry-Riddle Aeronautical University Extended Campus, the preceding feats need to be doable within a nine-week time frame or an equivalent thereof.

Suggested Sample Evaluative Vocabulary Models

The extrapolations of evaluative vocabulary designated as Model 1 and Model 2 were derived from online publications by Metfessel, N., Michael, W., and Kirsner, in "Instrumentation of Bloom's and Kratwohl's taxonomies for the writing of educational objectives," from a paper *Affective Domain* by N. Perrin, and J. Rueter, and from *Critical Thinking Skills: Building Blocks for Success* by F. Johnson. Such models provide quick access to sample evaluative vocabulary for the learning modalities, domains and critical thinking skills levels. Therefore, they have utility for enhancing the assessment of students' learning in a variety of circumstances.

SAMPLE EXTRAPOLATIONS OF EVALUATIVE VOCABULARY 1

MODEL 1

CRITICAL THINKING SKILLS LEVELS

MODALITIES	STEP 1 CORE	STEP 2 S & R	STEP 3 EXTRAP	STEP 4 UTILITY
VISUAL	identify locate observe read recognize watch	chart map review select Show visualize	apply change confirm differentiate examine identify interpret write	classify diagram graph illustrate modify research
AUDITORY	discuss listen question respond state	quote recall recite restate discuss	argue predict present project refute	categorize defend discuss significance persuade pro/con propose reject
KINESTHETIC	bend dance play sing skip	draw label mime paint Perform sketch	apply build create demonstrate sculpt	activate demonstrate perform conduct operate

It is neither expected nor desirable that one attempt to utilize all possibilities within the framework of each individual course. Rather it is the goal to bring the affective domain into balance with the cognitive and psychomotor growth and development of the students. This may, at times, involve very basic development of beliefs, attitudes, and values that will enhance, not impede, student progress. Instructors must sometimes counteract negative attitudes toward a specific course because the student or students perceive little or no value to themselves or they may simply not like a particular aspect of the course. Is this an affective student behavior that instructors need to address? It is, if one wishes to assist students to develop attitudes, beliefs, and values that have utility in the workplace, at home, and in their communities. Is it easy to do? Ongoing research indicates that it is not an easy task. Occasionally, the instructor must find ways to turn negative attitudes about course content into positive ones. Sometimes, an explanation of the rationale behind the course is sufficient. Sometimes, peer discussion of the pros and cons of the likely results of maintaining negative beliefs brings new insights into play. These may, in turn, foster the development of more positive attitudes and provide encouragement to reevaluate attitudes, beliefs, and values when situations and circumstances change. None of the preceding scenarios are scientific nor are they especially predictable. Hence, we have the appellation *affective domain* that is presented in Model 2.

SAMPLE EXTRAPOLATIONS OF EVALUATIVE VOCABULARY

MODEL 2

CRITICAL THINKING SKILLS LEVELS

DOMAINS	STEP 1 CORE	STEP 2 S & R	STEP 3 EXTRAP	STEP 4 UTILITY
COGNITIVE	define describe discuss explain recognize identify	analyze classify evaluate infer interpret paraphrase recall review summarize	assumptions connotation figures of speech interpret philosophies predict project satire speculate symbolism	analyze argue create draft evaluate examine options implement modify propose solve synthesize
PSYCHOMOTOR	(See kinesthetic modality on preceding chart)			
AFFECTIVE	acceptance acknowledge awareness follow listen willingness	consider define examine recall restate select specify	abstract balance compare contrast differentiate organize revise	adapt assist augment change defend formulate manage refute resist resolve revise subsidize support theorize

Options for the Evaluation of Students' Learning

The sample vocabulary suggestions in each of the critical thinking skills levels that are related to the cognitive domain involving the visual modality have appeared in some form in the framework for standardized testing that is often used for placement of students. In addition, such standardized tests are often used to compare student achievement with others in similar or like circumstances. Closer scrutiny of the auditory and kinesthetic modalities on standardized tests reveals that such questions are usually limited to listening to instructions and to filling in bubbles with a number 2 pencil. Unless the college/university levels of standardized testing have changed dramatically, only a few questions attempt to elicit responses related to the affective domain. As with commercially prepared tests and/or instructor prepared examinations, the preceding tests utilize Bloom's taxonomy in efforts to accurately measure results. Many vocabulary lists and charts that are rooted in Bloom's taxonomy are available from a variety of sources including those found online.

Commercially prepared examinations that are available from the textbook publishers provide one means of evaluating student progress but may not match each instructor's teaching style or points of emphasis for any given course. Furthermore, such tests or examinations may not address the affective domain and/or the modalities except as noted above or do so in a

superficial manner. A commercially prepared testing instrument developed by an expert in the field is or can be a valuable tool. If such is not the case, the validity of what is measured is brought into question. These and other concerns weigh into the decision by the instructor to create his/her own testing instrument.

What, then are the criteria to be observed when an instructor opts to create the assessment/evaluation tool to determine student progress and outcomes? Let us revisit the obvious for a moment. Well-written course objectives help define the nature and level of assessment. Generally, course objectives are or should be established within the framework of the course outline. The syllabus provides the opportunity for the instructor to determine the major focus of the course. This entails making decisions predicated on the needs of the students as identified in the course outline and on the time constraints that impact what to keep and what to leave out as well as the needs of the students themselves. Thus, begins the challenging but not insurmountable task of creating a valid examination. Planning and developing a satisfactory examination prior to meeting students is daunting and is done with little or no knowledge about the learning modalities of individual students. However, a variety of reasons including distance learning and other electronic media delivery of course content dictate that the planning and developing of a few of the examinations be done before classes open. Distance learning and interactive media delivery

of course content usually require that examinations be written prior to the beginning of the class. There are two valid reasons for doing this. One is that the instructor is likely to consider a variety of possibilities and prepare test items that access not only a variety of modalities but of domains as well in order to measure the outcome of student learning for each individual in a given course. The other reason is that students involved in distance learning or interactive media delivery of course content do not work at the same pace. Consequently, there often exists no fixed examination date such as is found in a traditional classroom. Furthermore, according to the previously cited New York Institute of Technology students retain more for longer periods of time that which has been learned if they participate in activities that engage them in "see, say, and do" paradigm. In a study "Effects of Anticipation of Tests on Delayed Retention in Learning" conducted by Haynie (1997), he found that anticipation alone without taking a test did not improve retention of learning. However, when the anticipation was followed by actually taking the test, retention did improve. As Haynie stated, "The conclusion here is that, in general, students do likely *study more earnestly* when they expect a test than if they do not, but maximum benefit in retention is gained only by having students anticipate and then actually take a test." The preceding rationale suggests that the writing of examinations prior to the opening day of a course may provide a greater sense of direction to both the students and the instructor. An added benefit for the instructor is that prior

preparation reduces the time needed to add or delete items from the test on an *as needed* basis.

Is it feasible to include the modalities and the domains of learning inclusive of the higher level critical thinking skills in the assessment of student learning as proposed by the models presented in this paper for most, if not all, courses offered at an institution of higher learning? Admittedly, it is not always practical or desirable in large introductory lecture classes with students sitting in a lecture hall that resembles an amphitheater. However, as a student moves into the higher levels of academia or participates in distance learning and/or interactive media, the number of students in any given course decreases to a size that makes such inclusion possible. This is especially true in an institution of higher learning such as Embry-Riddle Aeronautical University with its focus from its inception on meeting the educational needs of those engaged in the aircraft industry.

Flexibility of choice is an integral part of the models. The instructor determines which of the modalities and domains as well as which of the critical thinking skills levels (Johnson 2001) will be targeted for examination. It remains both the instructor's choice and responsibility to determine whether a specific assessment instrument will utilize multiple choice, short answer, short essay, long essay, research paper, an exhibit or demonstration, a performance, a project (group or individual), speech or some combination of the preceding. The schedule for the administration of the testing or assessment instruments is left to the discretion of the instructor within the time constraints of the course schedule. The decision

to choose to use textbook publisher prepared examinations or those of his/her own creation also rest with the instructor. The evaluation of students' growth and development may examine general knowledge and application or may be criterion-referenced (mastery tests) or some combination or the adaptation of the preceding possibilities. It remains in the instructor's domain to adapt both delivery of the material and method(s) of testing, assessment, or evaluation to the time frame within which such delivery occurs.

It is likely that the cognitive and psychomotor domains are examined more thoroughly at the basic or core knowledge and application levels throughout the educational process. Scrutiny of the psychomotor domain occurs more readily in such courses as physical education, the performing and industrial arts, etc. The higher levels of the cognitive domain are scrutinized in more detail as one enters into areas of specialization.

The most perplexing challenge that is inherent in the evaluation of student progress lies within the area of the affective domain. Evaluating the subjective area of attitudes, beliefs, and values is fraught with pitfalls. One such is that concepts may mean one thing to one person and the opposite or something very different to another. Nor are such concepts easily subjected to the scientific method of study that begins with a premise that may be proved or disproved. Another difficulty for the evaluator of the affective domain is that people may say or even believe one thing but do another. Various other difficulties are encountered when

attempting to assess the affective domain of student learning and behavior including much discussion and disagreement about what should be included in the domain itself. It is generally agreed that beliefs, attitudes, and values belong in this category with the focus on positive behaviors. Others argue that the terms are so broad as to be meaningless. However, the overriding question to be answered is to whom and by what standards do we bequeath the option of defining which beliefs, attitudes, and behaviors are positive? Are the beliefs, attitudes, and behaviors targeted for the course flexible enough to encourage creative and tolerant thinking as well as "by the book" should the situation warrant such a decision. Such questions could go on for some time. However, it is likely that the answers to the preceding questions are going to be influenced by the attitudes, values, and beliefs that relate to a multitude of issues.

As is evidenced by the strength of subjective concepts, it is necessary to determine the appropriate role of an institution of higher learning in defining these concepts. It is a given that the creation of such institutions was and is based in the underlying belief that those who are well educated fare better in life and in the world than those who are not. It is both implied and often stated that education changes the way one perceives and responds to situations in which one finds oneself. It is increasingly a given that if one wishes to avoid a dead-end job lacking challenges and opportunities for growth, one must be well educated. Even our television commercials espouse, and rightly so, that it is a

terrible thing to waste a mind. This is but one of many attempts to encourage youngsters to pursue an education beyond high school. Yet, definitive parameters are elusive and difficult to formulate in so diverse a society. One must be ever mindful of the ramifications of subjective attitudes, beliefs, and values that can be far reaching in unexpected and sometimes undesirable ways. Hippocrates, as translated by Francis Adams, states in his oath that reads in part, "I will follow that system of regimen which, according to my ability and judgment, I consider for the benefits of my patients, and abstain from whatever is deleterious and mischievous." Although his statement refers to the practice of medicine, one can by substituting the word *students* for *patients* establish baseline parameters for the testing, assessing, measuring, and/or evaluating the affective domain.

Model 2 (page 11) provides sample vocabulary lists that can be used to structure questions designed to elicit responses that reflect attitudes, beliefs, and behaviors. In addition, questions may be structured to engage the student in the examination of other and/or opposite attitudes, beliefs, and values. There does exist the possibility that students will attempt to answer according to what they perceive as "what the instructor wants to hear" rather than what they feel, believe, or do. Other methods of evaluating affective behaviors include self-evaluation scales predicated on the assumption that one can and does perceive in oneself that which others observe. A Likert-like scale with rankings from most preferred to least, from strongly agree to strongly disagree etc., may

be developed for this purpose. The problems with scales of this sort are twofold. The response may be an attempt to provide the answer that the student perceives to be the one the instructor is seeking. Or the response may be an honest attempt to report accurately but the student does not perceive the behavior in the same manner that others do. Psychological scales are used to identify maladaptive aspects of behaviors if the situation warrants and criteria for administering them are met. Such behaviors tend to fall within the affective domain. Generally, such measures are inappropriate for use in most college level settings unless one is entering a professional field demanding knowledge of administering and interpreting the results of such scales. Other highly specialized scales are used to determine types of leadership behavior as well as types of student or employee behaviors. However, the instruments that do tap into the affective domain tend to be administered so infrequently that they do not meet the criteria necessary to measure outcomes of student learning with respect to the affective domain.

Observations by the instructor of students' participation and interaction within group activities such as discussions, demonstrations, and research projects, etc. may reveal underlying attitudes, beliefs, and values. Such observations tend to be more subjective than other instruments used for measuring what students have learned. Putting aside momentarily the ethical arguments that were previously discussed, it is difficult but not impossible to assess such observations with a numerical point value. Let us return now to the ethical dilemma. It is real. It should be of

concern to us. It calls into question the tenets upon which our country is based as well as those upon which our institutions of higher learning rest. Yet, we recognize the need to discourage and penalize those who cheat on exams and those who plagiarize the work of others. Too much intrusion into the affective domain of an individual or society stifles the rights of those upon whom it is imposed. Too little development of the affective domain has the potential to become the *fast track* to decisions by individuals and societies that are detrimental to humanity. The challenge is to find some middle ground that is neither too restrictive nor so lacking in substance that it is rendered meaningless. It is also advisable to provide specific affective domain objectives to students in addition to those that address cognitive and performance objectives.

What is needed are methods to deliver instruction of course content in ways that also involve making judgments based on attitudes, beliefs, and values as well as the ramifications of those judgments upon the individual, friends, family, and the larger community. This is no small or insignificant task. In fact, it is a daunting one. Yet, the body of research indicates that many consider the task to be of the utmost importance, not only from the standpoint of the delivery and measurement of instruction but to the success of future endeavors by the students that we are charged to educate. Many leaders in business and industry are actively seeking to develop more positive beliefs, attitudes, and values in the workplace as evidenced by websites with postings related to the preceding.

How does one proceed? It has been suggested that the development of rubrics for use if observation is a part of the assessment of student learning of course content can be designed to control for subjectivity. Such a rubric might be used to observe interaction with peers, with group projects, etc. The development and use of rubrics to monitor such responses can identify the subjective elements and provide a basis for assigning point values to the task. Judith Slisz (2001) in "Examples of Interactions and Assessment Model for Online Courses Offered by Teikyo Post University" proposes the use of rubrics for the required online conferencing as well as for the individualized components of online courses. The rubrics differ in that the online conferencing component engages the affective domain through making supported value judgments about the work of others as well as about one's own. The rubric for course content has fewer affective and more cognitive domain aspects. Craig Mertler (2001) suggests that rubrics fall into categories, holistic and analytical. Furthermore, he has developed templates, charts, and a design plan to assist with the implementation of the use of rubrics to evaluate the outcomes of student learning.

For many courses such as the humanities, social sciences, communications psychology, and others, it is possible to construct questions and activities that elicit value judgments supported by evidence from both assigned readings and research. The logical and insightful development of an answer can provide further insights into aspects of the affective domain. Open-ended questions can be developed for many kinds of

course content thereby creating opportunities for students to examine the rationale underlying their beliefs, attitudes and behaviors as well as those of the authors or developers of texts, theories, proposals, etc. Role playing, speech making, cooperative learning projects, panel discussions in a quasi-debate format and cognitive questions that elicit opposite or differing opinions, beliefs and opinions may be used to integrate the learning modalities with the learning domains. The sequencing of affective objectives accord to Krathwohl's hierarchy of receiving, responding, valuing, organizing, and characterizing was suggested by Barbara Martin in 1989.

Time constraints inherent in any academic offering often leave time for little more than a cursory attempt to involve the affective domain in most cognitive domain oriented courses. This is especially true when working with nine-week sessions or variations thereof. However, a longitudinal study by Richard M. Felder (1995) and Felder, Felder, and Dietz (1998) in an ongoing five course program for engineering students suggests that scores improved and that retention of learning was better for those in the experimental cooperative learning group than for those in the control group. Attitudes toward cooperative learning became more positive as students became more accustomed to working as members of groups as indicated by the following:

In the semester following the experimental course sequence, the students were asked to evaluate the sequence retrospectively. Of the 67 seniors responding, 92% rated the experimental courses more instructive than their

other chemical engineering courses, 8% rated them equally instructive, and, none rated them less instructive. Ninety-eight percent rated group homework helpful and 2% rated it not helpful, and 78% rated in-class group work helpful and 22% rated it not helpful. (Felder 1995)

Felder noted that the Hawthorne effect could have affected the results of his study. However, his findings do suggest that attitudes, beliefs, and values affect students' interactions, performance, and retention of learning in the classroom setting. His findings further reinforce the rationale for seeking ways to broaden the scope of the measurement of the outcomes of student learning to include the affective domain. It should be noted that cooperative efforts in the workplace have been linked by some to greater productivity and improved safety in that environment.

Summary

Admittedly, measuring or assessing the affective domain to determine what has been learned and what has been retained such that it may be used is oftentimes an elusive task. It is subjective, not easily measured, and is sometimes treated as if there is something inherently wrong with planned development of values, beliefs, and attitudes. Yet, the body of literature published during the recent and not so recent past indicates that education especially at the college/university level is predicated upon changing one's beliefs, attitudes, and values to ensure realizing one's potential for intellectual growth and development. The literature related

to testing indicates that the cognitive domain with its left brain thinking and predominately visual modality is the most thoroughly tested aspect student learning. Assessment of the psychomotor domain relies heavily on performance (kinesthetic modality and right brain activity) in certain performing and mechanical arts and physical education arenas. However, it is possible to include activities, projects, etc. that will involve some facet of the domains, modalities, and critical thinking skills in each of the courses that rely heavily on the more traditional cognitive level. The proposed models with sample vocabulary provide a brief checklist to enhance the inclusion of a variety of levels of questions, activities, projects, etc. covering the learning modalities and the recognized domains involved in learning. It is not all-inclusive nor should it be. It is the prerogative of each instructor to make the final decision regarding the format as well as the domains and modalities to be tested.

There is no scarcity of literature about any of the preceding topics. However, it is my belief that the Johnson's Sample Extrapolations of Evaluative Vocabulary Models 1 and 2 represent a more comprehensive approach to directly linking the domains of students' learning as well as the learning modalities to critical thinking skills levels in a user friendly manner. The models are designed as thought provoking and time saving guides to enhance the formulation of questions in appropriate domains, modalities, and critical thinking skills levels at the instructor's discretion in a specific field. It has application in most, if not

all, courses including online and interactive delivery of course content.

In addition to the preceding concepts, the key to successful measurement of students' learning rests within the framework of well-written course outlines and well-written objectives. It is upon such a foundation that the ability to design and implement effective measurements of student learning in both academic and technical and/or job related learning situations resides. But that research will have to wait for another forum.

Sources

- Affective domain value development, personal development, & esthetic development. (2001). *Affective domain, classification of learning skills*. Retrieved on 03-09-02 from <http://pcrest.com/affective.htm>
- Bandura, A. (1986). Self-efficacy beliefs in human functioning. [Electronic version]. *Social foundations of thought and action*. Retrieved on 05-01-02 from <http://www.emory.edu/EDUCATION/mfp/effpassages.html>
- Bartell, et al. (1988). Code of fair testing practices in education. Washington, D.C. *Joint committee on testing practices*. Retrieved on 02-20-02 from <http://ericae.net/code.txt>
- Bloom's taxonomy. Retrieved on 04-09-02 from <http://www.nyiteez.org/EDIN777/bloom.htm>
- Bloom's taxonomy of educational objectives. *University of Queensland assessment site: Bloom's taxonomy of educational objectives*. Retrieved on 03-09-02 from <http://www.tedi.up.edu.au/Assess/Assessment/bloomtax.html>
- Bond, L. A. (1996). Norm- and criterion-referenced testing. *Practical assessment, research and evaluation*. Retrieved on 02-20-02 from <http://ericae.net/pare/getvn.asp?v=5&n=2> (ISSN 1531-7714)
- Boultinghouse, A. Learning styles: accommodating student needs. *Changing the way we teach*. Retrieved from <http://.aea1.k12.ia.us/mentoring/learnteach91201.pdf>
- Cassady, J. C. (2001). The stability of undergraduate students' cognitive test anxiety levels. *Practical assessment, research & evaluation*, 7(20). Retrieved on 02-20-02 from <http://ericae.net/pare/getvn.asp?v=7&n=20>
- The CHES exam. (2002). The National Commission for Health Education Credentialing, Inc. Retrieved on 02-20-02 from http://www.nchec.org/ches_exam.htm
- Cognitive domain information processing, critical thinking, problem solving, & research. (2001). *Cognitive domain, classification of learning skills*. Retrieved on 03-09-02 from <http://pcrest.com/cognitive.htm>
- Cognitive theories of learning. *NSW HSC ONLINE - Professional development*. Retrieved on 03-06-02 from http://hsc.csu.edu.au/pro_dev/teaching_online/how_we_learn/cognitive.html
- Constructivist theory. *TIP: Theories*. Retrieved on 05-01-02 from <http://tip.psychology.org/bruner.html>
- Criterion referenced instruction. *TIP: Theories*. Retrieved on 05-01-02 from <http://tip.psychology.org/mager.html>
- Conditions of learning. *TIP: Theories*. Retrieved on 03-06-02 from <http://tip.psychology.org/gagne.html>
- Critical and creative thinking - Bloom's taxonomy. (2001). Retrieved on 02-24-02 from <http://www.eduscapes.com/tap/topic69.htm>
- Daniels, M. Maslow's concept of self-actualization. Retrieved on 05-01-02 from <http://www.mdani.demon.co.uk/archive/MDMaslow.htm>

Edwards, B. (1989). *Drawing on the right side of the brain*. Los Angeles. Jeremy P. Tarcher, Inc.

Felder, R. M. (1995). A longitudinal study of engineering student performance and retention. IV. Instruction methods and student responses to them. *Journal of engineering education*, 84(4), 361-367. Retrieved on 04-23-02 from <http://www2.ncsu.edu/unity/lockers/users/ff/felder/public/Papers/long4.html>

Felder, R., Felder, G., & Dietz, E. (1998). A longitudinal study of engineering student performance and retention. V. Comparisons with traditionally-taught students. *Journal of engineering education*, 87(4), 469-480. Retrieved on 04-23-02 from <http://www2.ncsu.edu/unity/lockers/users/ff/felder/public/Papers/long5.html>

Find your learning style. *West Chester University learning assistance and resource center*. Retrieved on 04-16-02 from <http://www.wcupa.edu/academics/cae.tut/Modality.htm>

Hayes, N. The psychology of interactives. *British Interactive group - research - psychology of interactives*. Retrieved on 03--6-02 from http://www.big.uk.com/research/psych_interactives.htm

Haynie, W. J. (1997). Effects of anticipation of tests on delayed retention of learning. *Journal of technology*, 9(1). Retrieved on 04-23-02 from <http://scholar.lib.vt.edu/ejournals/JTE/v9n1/haynie.html>

Heylighen, F. (1992). A cognitive-systemic reconstruction of Maslow's theory of self-actualization. [Electronic version]. *Behavioral science* 37, 39-57. Retrieved on 05-14-02 from <http://pespmc1.vub.as.be/Papers/Maslow.pdf>

Hippocrates. (400 B.C.E.). *The Oath*. (F. Adams, Trans.). Retrieved on 05-29-02 from <http://classics.mit.edu/Hippocrates/hippooath.html>

Hooper, C. (1998). *Left vs. right which side are you on? Practicing college study skills: strategies for success*. Houghton Mifflin Company. Retrieved on 05-31-02 from <http://www.mtsu.edu/~devstud/advisor/LRBrain.html>

How we learn. Retrieved on 04-16-02 from <http://www.nyiteez.org/EDIN777/howlearn.htm>

Huitt, W. (1999). The affective domain. *Educational psychology interactive: the affective system*. Retrieved on 05-11-02 from <http://chiron.valdosta.edu/whuitt/colaffsys.html>

Huitt, W. (1999). Conation as an important factor of mind. Retrieved on 05-01-02 from <http://chiron.valdosta.edu/whuitt/col/regsys/conation.htm>

Huitt, W. & Hummel, J. (1997). Observational (social) learning. *Educational psychology interactive: observational learning*. Retrieved on 05-01-02 from <http://chiron.valdosta.edu/whuitt/col/socog/soclm.htm>

Huitt, W. (1997). Socioemotional development. *Educational psychology interactive: socioemotional development*. Retrieved on 05-31-02 from <http://chiron.valdosta.edu/whuitt/col/affsys/erickson.html>

Huitt, W. (1997). Values. *Educational psychology interactive*. Retrieved from <http://chiron.valdosta.edu/whuitt/col/affsys/values.htm>

Internal processes of learning. *Learning processes*. Retrieved on 04-09-02 from <http://www.nyiteez.org/EDIN777/learningprocesses.htm>

Measuring Outcomes Of Students' Learning

- Jansak, K. E. Bloom's taxonomy in the cognitive domain for indexing levels of abstraction required in assessments & assignments. *The Ohio learning network. Instructional design for web-intensive courses workshop*. Retrieved from http://www.edison.cc.oh.us/online/instructional_design.htm
- Johnson, D. Teaching styles: what are they? And why do I need them? *Music staff teacher lounge article 27*. Retrieved on 05-31-02 from <http://www.musicstaff.com/lounge/article27.asp>
- Johnson, F. (2001). Critical thinking skills: building blocks for success. E. Wheeler (ed.), *College of Career Education Ninth Annual Symposium on Teaching Effectiveness*. Daytona Beach, FL: Embry-Riddle Aeronautical University.
- Krathwohl, D. R. (1964). Descriptions of the major categories in the affective domain. [P. Schneller, Electronic posting]. Retrieved on 05-29-02 from <http://www.muc.edu/~schnelp/Krathwohl's%20Taxonom.html>
- Learning modalities. Retrieved on 05-29-02 from <http://its.foxvalley.tec.wi.us/iss/curric-assessment/modaliti.htm>
- Learning pyramid. [Tehama County Department of Education, Electronic posting]. Bethel, MA: National Training Laboratories. Retrieved from <http://www.tcde.tehama.k12.ca.us/pyramid.pdf>
- Learning retention. (2001). *FNF: learning retention (How much do we remember?)*. Retrieved on 05-28-02 from http://www.hi.is/~joner/eaps/cs_reten.htm
- Learning styles and the 4MAT system: a cycle of learning. [Electronic version]. B. McCarthy. *4MAT in action: creative lesson plans for teaching to learning styles with right/left mode techniques*. Retrieved on 05-31-02 from <http://www.volcano.und.nodak.edu/vwdocs/msh/lc/is/4mat.htm>
- The Likert scale. *Likert scale - measurement of attitudes*. Retrieved on -5-11-02 from <http://www.cultsock.ndirect.co.uk/MUHome/cshtml/psy/likert.html>
- Livingston, J. A. (1997). Metacognition: an overview. Retrieved on 04-24-02 from <http://www.gse.buffalo.edu/FAS/Shuell/CEP564/Metacog.htm>
- Martin, B. L. (1989). A checklist for designing instruction in the affective domain. Retrieved on 03-06-02 from <http://plaza.v-wave.com/kegj/mar.htm>
- Mcleod, S. H. (1991). The affective domain and the writing process: working definitions. Retrieved on 03-06-02 from <http://jac.gsu.edu/jac/11.1/Articles/6.htm>
- McNabb, J. & Mills, R. (1995). Tech prep and the development of personal qualities: defining the affective domain. *Summer95 115(4)*, 589(4p). Retrieved from <http://www.epcc.edu/Programs/Americana/Critical/CtarticleMcNabb.htm>
- Mertler, C.A. (2001). Designing rubrics for your classroom. *Practical assessment, research & evaluation 7(25)*. Retrieved on 02-20-02 from <http://ericae.net/pare/getvn.asp?v=7&n=25>
- Metfessel, N., Michael, W. & Kirsner, D. Instrumentation of Bloom's taxonomy for the cognitive and affective domains. *Instrumentation of Bloom's and Krathwohl's taxonomies for the writing of educational objectives*. Kibler, R. et al. *Behavioural objectives and instruction*. Boston, Allyn & Bacon. Retrieved on 05-11-02 from <http://www.unesco.org/webworld/ramp/html/r8810e/r8810e0e.htm>
- Moskal, B. M. Scoring rubrics: what, when and how? *Practical assessment, research and evaluation*. Retrieved on 02-20-02 from <http://ericae.net/pare/getvn.asp?v=7&n=3> (ISSN 1531-7714)

- Parajas, F. (1996). Assessing self-efficacy beliefs and academic outcomes: the case for specificity and correspondence. *Proceedings of the American Educational Research Association*, New York. Retrieved on 05-01-02 from <http://www.emory.edu/EDUCATION/mfp/aera2.html>
- Parajas, F. (2000). Overview of social cognitive theory and of self-efficacy. Retrieved 05-01-02 from <http://www.emory.edu/EDUCATION/mfp/eff.html>
- Perrin, N. & Ruter, J. Some classroom techniques useful in assessing different affective domains. *Affective domain @ study mate*. Retrieved on 04-09-02 from <http://www.nyiteez.org/EDIN777/affective.html>
- The problem of scoring. Retrieved on 03-09-02 from <http://www.criticalthinking.org/University/pseudo/8Problem.html>
- Psychomotor domain. (2001). *Psychomotor domain, classification of learning skills*. Retrieved on 03-09-02 from <http://pcrest.com/pmotor.htm>
- Reeves, M. F. An application of Bloom's taxonomy to the teaching of business ethics. *Journal of business ethics* 9. (609-616J). Netherlands, Kluwer Academic Publishers.
- Reigeluth, C. M. What is the new paradigm of instructional theory. *Proceedings from ITFORUM paper 17*. Indiana University. Retrieved on 03-06-02 from <http://itech1.coe.uga.edu/itforum/paper17/paper17.html>
- Reisman, F. K. (1982). Diagnosing the affective domain. *A guide to the diagnostic teaching of arithmetic*. Columbus, OH, Charles E. Merrill. Retrieved on 05-01-02 from <http://www.drexel.edu/academics/teachered/hdl/manual/maslow.htm>
- Riddle, E. M. & Dabbagh, N. (1999). Lev Vygotsky's social development theory. Retrieved on 05-01-02 from <http://chd.gse.gmu.edu/immersion/knowledgebase/theorists/constructivism/vygotsky.htm>
- Right brain vs. left brain. (1998-2001). *On Purpose Associates*. Retrieved on 05-31-02 from http://www.funderstanding.com/right_left_brain.cfm
- Roeber, E. D. (1996). Guidelines for the development and management of performance assessments. *Practical assessment, research, & evaluation*, 5(7). <http://ericae.net/pare/getvn.asp?v=5&n=7>
- SCAN/TEJAS affective assessment sample. *Test of employability: job aptitudes & skills*. Retrieved on 03--09-02 from <http://www.ric.dcccd.edu/annex/id/affectiv.htm>
- Shadish, W. Some evaluation questions. *Practical assessment, research & evaluation*, 6(3). Retrieved on 02-20-02 from <http://ericae.net/pare/getvn.asp?v=6&n=3>
- Simon, M. & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. *Practical assessment, research & evaluation*, 7(18). Retrieved on 02-20-02 from <http://ericae.net/pare/getvn.asp?v=7&n=18>
- Slisz, J. (2001). Examples of interactions and assessment model from online courses offered Teikyo Post University. *Proceedings from the NEASC conference on Assessing the impact of different modes of interaction of student learning in online classes*. Retrieved on 05-28-02 from http://www.neasc.org/cihe/modes_interaction.PDF
- Social cognition. (1998). *Educational psychology interactive: social cognition*. Retrieved on 05-01-02 from <http://chiron.valdosta.edu/whuitt/col/soccog.html>

Measuring Outcomes Of Students' Learning

Social constructivism. Retrieved on 05-01-02 from <http://pdts.uh.edu/~srmehall/theory/social.htm>

Social domain. (2001). *Social domain, classification of learning skills*. Retrieved on 03-09-02 from <http://pcrest.com/social.htm>

Social learning theory. *TIP: Theories*. Retrieved on 04-23-02 from <http://tip.psychology.org/bandura.html>

Subsumption theory. *TIP: Theories*. Retrieved on 05-01-02 from <http://tippsychology.org/ausubel.html>

Steps in assessment of student outcomes. Andrews University Assessment Office. Retrieved on 03-09-02 from <http://www.andrews.edu/assessment/steps.htm>

Taylor, S. M. Cooperative learning in distance education. *IHETS: Cooperative learning in distance education*. Retrieved on 05-29-02 from http://www.ihets.org/learntech/distance_ed/fdpapers/1997/taylor.html

What are promising ways to assess student learning. (Spring, 1996). *Improving America's School: A newsletter on issues ins school reform*. Retrieved on 02-20-02 from <http://www.ed.gov/pubs/ASA/newsletters/assess/pt3.html>

What issues affect Likert-scale questionnaire formats? Questions and answers about language testing statistics: Likert-scale questionnaires. *Shiken: JALT testing & evaluation SIG newsletter*. Retrieved from http://www.jalt.org/test/bro_7.htm

What the research says about student assessment. (Spring, 1996). *Improving America's school: A newsletter on issues in school reform*. Retrieved on 02-20-02 from <http://www.ed.gov/pubs/IASA/newsletters/assess/pt4.html>

WIDS affective domain. Retrieved on 03-06-02 from <http://its.foxvalley.tec.wi.us/iss/curric-assessment/affectivedomain.html>

**Peer Evaluations, and Team Learning In
Undergraduate and Graduate Education**

Bruce A. Rothwell

Abstract

This paper reviewed the available literature concerning the use of student peer evaluations and team learning in undergraduate and graduate education. The literature clearly demonstrates that the simple act of how an instructor assigns groups has a significant impact on the grades their students receive. Also clear from the literature is that training should precede the use of peer evaluations in any classroom setting. Further, student assigned grades become more accurate with training and practice. Although there are several potential problems with the use student peer evaluations and team learning, there are also very pronounced student benefits. Generally, students have been shown to improve their learning, retention, and overall grades when group projects and peer evaluations are correctly used.

Peer Evaluations

"Peer evaluation is the process of having the members of a group judge fellow members on specified traits, behaviors, and achievements" (Sherrard, Raafat, & Weaver, 1994, p. 43). Keith Topping (1998) accomplished an extensive review of literature covering peer assessments at higher education institutions. He reviewed 109 different articles published between 1980 and 1996. Of the 109 articles, 67 were quantitative and presented data the authors reviewed had collected during their research on peer evaluations. The remaining articles that were reviewed were descriptive in nature. His conclusion about the literature prior to 1996 was significant. Although Topping (1998) found a large number of studies had been accomplished, the procedures used and varied disciplines involved in the research made a definitive decision on the effectiveness of peer evaluations impossible.

Even though an overall definitive finding on the value of peer evaluations was not possible, specific aspects of the literature made it possible for Topping (1998) to formulate several meaningful and valid conclusions. The sheer number of different academic disciplines represented by his research lead to the conclusion that peer evaluations may be applicable in all academic areas. Nearly all the studies reviewed indicated that the instructors had

mandated the peer evaluations instead of the students initiating them. Although the process where instructors who pushed the peer evaluation system on their students was seen as a possible problem, the process of involving students in peer evaluations resulted in students getting more involved in the subject matter. Peer evaluations created a sense of student ownership and required the students to take responsibility for the evaluations they provided (Topping, 1998).

Although a definition of peer evaluations is easily definable and generally agreeable to most, implementation strategies and the value placed on them differs greatly. Some instructors see them as merely beauty contests while others have found them to be an important part of the learning process. In part, those who have used and liked peer evaluations found that they lent a unique perspective that the instructor could not possess. Specifically, peer evaluations come from those individuals closest to the individual being rated and who are also the people who have the most contact with them (Sherrard, Raafat, & Weaver, 1994).

When instructors assign people to team projects, the professor must understand that there will invariably be some students who will attempt to ride the coattails of the rest of the group. These free-riding individuals do not participate as

much as the remainder of the group and in some cases may not participate at all (Dyrud, 2001). Left alone, student groups may or may not effectively solve the various types of problems they can encounter. In many cases, the instructor is not even aware of discontent within the student teams until very near the time for the project to be delivered for a grade. Generally, near the end of term, the students are frustrated and group interaction may be well beyond the point of simple disagreement. One way to alleviate some of the problems found within group assignments is for the instructor to use a system of peer reviews within the group (Dyrud, 2001).

Terry Gatfield (1999) suggests that groups permeate the working environment because a group's many different talents allow it to accomplish what no individual person could do alone. A natural outgrowth of the group-working environment was to bring the use of groups into the classroom where they can improve student learning. The use of groups in a class is reported to have many advantages including the finding that students will learn first-hand how to interact within a group and become better prepared to take their place in a work environment where groups have become prevalent. Following logically to the next level, the use of groups in the class created a need for instructors to understand the contribution of each individual student. To solve this need, a system of peer

assessments was established and used in various classrooms.

Most instructors found that creating teams in the classroom was the easy part. However, once the groups were created, one of the main problems then facing the instructor was how to grade the efforts of a group project. The final grade for an individual project is a relatively easy process of reviewing the work and assigning an appropriate grade. Because only one person worked on the assignment, the grade received should be a fair estimate of the effort the student put into the assignment. However, when more than one person has contributed to the completion of a project, there are little formal means for the instructor to know how much of a contribution was made by each individual. Therefore, one or more of the students may receive a free ride while others in the group did more than their fair share. Conversely, it is possible that a poor final project may have been submitted from a team comprised of several members who did little and one student who tried exceptionally hard. Any time an instructor assigns the same grade to all members of a team, it is probable he or she will penalize some and reward others based on factors outside of the students' control (Maranto & Gresham (1998).

Many studies have found significant benefits derived from using peer evaluations. If used correctly, peer evaluations will have a positive impact. It is possible for the peer evaluation process to

improve student grades and their overall involvement in course activities (Topping, 1998). Additionally, when students know that their efforts are going to be judged by fellow classmates, who may also be their friends, the students may be motivated to try harder than if they are only being evaluated by the instructor (Hite, 1996).

One important way of assisting the instructor in the assignment of a fair grade for group assignments is to have the members of the team tell the instructor how much of an effort each member of the team made to the project. The process of students reporting on other students is referred to as peer evaluations. Sherrard, Raafat, and Weaver (1994) performed an analysis of students who received peer evaluations and an instructor assessment for the same in-class presentation. The peer evaluations received determined 20 percent of the students' final grade in the course. In addition to the peer evaluations, each student rated their own performance on two in-class presentations. The researchers found that the self-evaluations the students provided were very accurate in comparison to evaluation scores they received from their fellow students. The researchers also found that only gender had an impact on the evaluations students received with females rating individuals higher than males. Finally, each student accomplished an end of course critique where they were asked specific questions about the peer evaluation process. The

majority of students reported that the peer evaluations were a valuable means of providing feedback and indicated that they had actually learned from the evaluations. However, the majority of students also felt that their peers had evaluated them on factors other than just the presentation on which they should have been assessed (Sherrard, Raafat, & Weaver, 1994).

In a different study, a researcher had students predict how fellow students would do on a final examination based on their perceptions of them during the term. Students had only in-class discussions and seminar presentations on which to base their predictions. Although the study only involved 75 undergraduates in their third year, the results reported were noteworthy. The researcher found that the students' peers accurately predicted how the students would perform on the final examination. The results were even more accurate of a predictor when the one being evaluated and the one doing the evaluation were most similar in ability and performance during the class. In other words, the closer the rater and the one rated were in the final class ranking, the closer the prediction was to where they would finish (Orpen, 1994).

In a study of 59 graduate students, Kelmar (1993) found that the students fairly accurately assessed the performance of their peers on an in-class presentation of an outside reading assignment. Unlike many studies, this research used the results of the peer assigned evaluations to determine 15

percent of the students' final grade. The students were not aware that their instructor would also be grading the presentation and would compare the student provided evaluations to his. Before the assessment, the students were trained in the dynamics of peer evaluations. Kelmar (1993) found that the student assessments were on average, significantly higher than his were. Further, the instructor found six of the presentations to be unacceptable graduate level work while none of the students assigned a failing grade to any of their peers. Although the students' evaluations were higher on average, they were all tightly centered about the mean, whereas the instructor's ratings had a much larger variation. The author attributed the differences in scoring to three main reasons. The students were seen as more sympathetic to their peers, the students were not previously experienced in assigning grades, and the students did not have the same opportunities to converse with other raters as the instructor had to discuss with his peers (Kelmar, 1993).

Marilyn Dyrud (2001) used a series of peer reviews designed to improve group interaction and report the contribution level of each member within the group. To this end, she used three mandatory peer reviews. The first two were open evaluations and were designed to let the group solve their own problems, promote group interaction, and eliminate undesirable group behavior as early as possible in the group project. The final peer review was

only seen by the instructor and was the one that she used to assign a portion of the final grade to the various group members. Group work was seen as essential by the researcher to prepare her students for what they will experience in the majority of work settings (Dyrud, 2001).

Gatfield (1999) studied 261 undergraduate students and found that the students were very satisfied with the system of peer evaluations used in his class. Further, he found that students who had previously been in an actual working environment had a higher level of group satisfaction than those students who had no previous work experience did. Lastly, he found there was no statistical difference between males and females in acceptance of groups and peer evaluations in the classroom. The one caution that the author presented was that groups and peer evaluations might not be appropriate for first year undergraduates because they may not possess the necessary prerequisites to handle group dynamics.

Persons (1998) determined whether or not factors that existed prior to the start of her accounting classes and those factors acquired during the accounting class had an impact on the peer evaluation students received. She called the factors that existed prior to being placed in a group learning setting ex-ante factors and those that were acquired while in the group ex-post factors. The ex-ante factors included gender, race, GPA and prior accounting background. The

ex-post factors included test scores, participation in class, and grades for group assignments. The researcher, who tried to distribute the various ex-ante factors as equally as possible, assigned the groups. All peer evaluations were conducted at the end of the course. Persons (1998) found that gender, race, and previous accounting knowledge had no impact on the peer evaluations received. However, the students' previous GPA and declaration of an accounting major were positively related to the peer evaluations they earned. When looking at the factors that students acquired since the start of the course, only their participation and group homework grades were positively related to their peer evaluation scores. The major findings from this study were that future instructor assigned learning teams should be equally filled based on previous GPAs and declaration of a relevant major (Persons, 1998).

Many instructors have developed creative ways around the problem of assigning group and individual grades for a group assignment. Most of these means center around some form of peer evaluations. One such peer evaluation system is known as the Knickrehm Method. Within the Knickrehm Method, the instructor assigns a group grade and members of the group assign a specific number of points to all other members of the group based on the students' contribution to the group project. Each member of the group is allocated a

specific number of points that they award to anyone in their group other than themselves (Maranto & Gresham, 1998).

Typically, the Knickrehm style has point ranges from zero points, awarded to members who did not contribute, to a maximum of four points, for members who did most of the work. Members are allocated a specific number of points sufficient to award everyone on their team with two points and still have one point left over to award to the best performer. In this manner, the majority of the members are awarded two points and one person receives three. The description for award of two points is that the person contributed their fair share. Although the awarding of two points is the norm, each person can give any other member zero or one point, leaving a higher possible score for someone else (or more than one other person) of their choosing. Because the instructor limits the percentage of the final grade that the peer evaluations impact, the points assigned by peers within the groups can only change an individual's grade on the margins and generally never more than ten percent (Maranto & Gresham, 1998).

Maranto and Gresham (1998) reviewed teams at two different universities to determine the impact of using the Knickrehm Method for peer evaluations. The first author found high student satisfaction with the method and in over 40 groups comprised of between four and seven members had only one student

complaint concerning the evaluations. Findings from the first author show that those scores that were reduced were lowered from four to eight percent and scores that were increased were improved from eight to 16 percent. Seldom did these changes have an impact on the final grade assigned. The second author was working with much larger class sizes and was met with high levels of student dissatisfaction. Although there were more students' complaints, about ten percent of the grades were lowered and 38 percent were increased through the peer evaluations. Part of the reason for the higher student dissatisfaction in the second scenario was that the instructor was at a university using the plus-minus grading system. Therefore, any movement in these peer evaluations away from the Knickrehm normal score of two had an increased potential for changing the students' final grade. The overall conclusion from this study was there might be more appropriate settings than others for using peer evaluations (Maranto & Gresham, 1998).

Peer evaluations can be effectively used on several different assignment types including oral presentations, group projects, and individual writing assignments. "An important feature of most WAC [Writing Across the Curriculum] programs is the use of peer review and peer grading; that is, students' evaluation of the writing efforts of their peers" (Kerr, Park, & Domazlicky, 1995, p. 357). One research study reviewed

by Kerr, Park, and Domazlicky (1995) found that when undergraduate business communication instructors used peer evaluations of writing assignments, the majority (82 percent) had their students review peer papers prior to the student turning in the paper for a grade. The remaining 18 percent had fellow students assign some form of an assessment of the paper turned-in that was used in the determination of a final grade. There are several positive reasons for having each student review their peers' papers. One of the main advantages is that by reading other papers, the students are exposing themselves to additional information concerning the course topic. The additional exposure to course materials helps retention and understanding of key points. Further, by knowing that their peers will see their work, the student will have an increased incentive to do their best because they do not want to look bad in front of their friends and peers. Closely related to this benefit, the students, because of the closer relationship they have with their peers, might place more weight on the opinions of fellow classmates than in their instructor's evaluation.

Although most studies involving peer evaluations involve performance within a group setting, one study tracked the evaluations of sophomore and junior undergraduates' evaluations of two individual writing assignments accomplished over the course of a single term. Although

the results showed that the students' evaluations consistently exceeded that of the instructor's evaluations; the differences were most pronounced in the scoring of the first paper. By the second paper of the term, peer evaluations decreased while the instructors' evaluations had increased. The increased instructor evaluations were seen as caused by the students' improved writing abilities resulting from the accomplishment of the first paper and the additional learning that took place from the review of their peers' first paper. One of the main findings of this study was that by learning how to evaluate others, the students also learned how to more critically evaluate their own efforts. The decrease in the students' evaluations of their peers was seen as being caused by learning what to look for on the first paper. Therefore, the students' ability to evaluate improved with practice (Kerr, Park, & Domazlicky, 1995).

Potential Problems with Peer Evaluations

Peer evaluations generally create significant initial anxiety in both the instructor and the student (Topping, 1998). Therefore, before attempting to implement a team learning approach, the professor must understand that there is much more to the process than simply assigning students a group project. The instructor has to provide the foundation and continuously follow-up to make sure the team learning approach is successful (Hite, 1996).

In recent years there has been an increasing trend to involve students in the evaluation process. Generally, these models fall into the two general areas of either self-evaluation or peer evaluation. Two of the biggest problems with the growing use of peer evaluations are that the students doing the evaluations may not be qualified raters and also may not be impartial in assigning their evaluations. Simply stated, some students may rate others of their same sex or race differently than those of the other sex or races (Ghorpade & Lackritz, 2001).

Ghorpade and Lackritz (2001) reviewed the performance of 221 senior level undergraduate students in human resource management to determine if there were differences in the way fellow students rated their peers based on sex and race when compared to previous class participation. The assignment that was peer evaluated was an in-class presentation provided by a team of three to five students. Each student received a peer evaluation on both an individual and team basis. The authors also evaluated each student. They then compared the amount of student participation they observed against their evaluations and against the evaluations provided by the students based on sex and race differences. Ghorpade and Lackritz (2001) also had each student self-evaluate their participation. The researchers then compared the students' self-evaluation with the authors' assessments of the student. In

85 percent of the cases, the students' self-evaluation matched that of the researchers. The researchers' findings also clearly showed that the more students participated in their classes, the higher that students' peer evaluations were. In fact, "[t]he single most significant influence behind the ratings process was frequency of participation in classroom discussions by the presenters" (Ghorpade & Lackritz, 2001, p. 279). Also interesting, was that there was no difference in the rating received or given based on the sex of the individual. In other words, men and women rated each other the same. Although there were no differences found for gender, the same was not true for differences in races, with African Americans receiving the lowest scores. Further, the Asian Americans did significantly better than any other group. Strangely enough, the highest ratings for whites came from the Asian and African American groups. Although not an initial consideration of their study, Ghorpade and Lackritz (2001) found that age also had an impact on the ratings students received. The older the student doing the presentation, the higher the rating they received. The overall finding of their research was that peer evaluations should not be used as the sole means of grade assignment.

Lejk, Wyvill, and Farrow (1999) report that a vast majority of United Kingdom professors (95 percent) reported using some form of group assessment at least once. Further, they indicated that the majority also

felt an "...uncertainty about the reliability of group assessment, especially when all group members are allocated the same grade" (p. 5). What makes the Lejk, Wyvill, and Farrow (1999) study significant was that their research spanned four years and tracked 729 university students who they had assigned to work within groups. About half of the 729 students were assigned to groups based on their performance on two tests they took before being assigned into their groups. The researchers separated the test scores into three classifications; low, medium, and high and assigned students into their groups using these three classifications. One half of their students were teamed with other students within their same classification. Specifically, students with high scores were teamed with other students with high scores while students with low scores were assigned to teams with other low scoring students. The other half of their students were assigned within groups of mixed performance results on the first two tests. These teams were comprised of three students with one from each of the high, medium, and low performance categories.

The results the researchers found were striking. Students who had done the best on the first two tests averaged 11 percent lower marks if they were assigned to mixed groups than those who had previously done well and were assigned with others who had also done well. Students who had done poorly on the first two tests scored an average of 12 percent higher when they

were assigned to mixed groups than those who were assigned with others who had not initially done well. The implications of these findings are clear, "[t]he method by which a group is formed seems to have an effect upon the performance of the group" (Lejk, Wyvill, & Farrow, 1999, p. 13). The act of simply assigning students to groups has a tremendous impact on the students' final grade and more importantly the amount of learning that they take away from a class. Randomly assigning students to groups would tend to push all scores toward the middle. While assigning by ability will hurt either the poor or high performers depending on whether they were assigned to mixed groups or groups of equal abilities.

Cheng and Warren (1999) determined if there were differences in the scores provided by students and instructors for first year electrical engineering students. The researchers first trained their students on what they should look for when evaluating their peers. Separate assessments were provided from both the instructors and the students for each written and oral assignment. The researchers found a significant difference between the grades provided by the instructors and the students. The students consistently grouped scores together with little variation in grades. The instructor assigned grades had a greater variation with a larger range of scores. Although there was an initial significant disagreement between the student and instructor assigned grades, as

the students graded more assignments, the differences started to shrink. The authors saw the narrowing of the differences between instructor and student assigned grades as occurring naturally as the students gained experience in grading (Cheng & Warren, 1999).

Cheng and Warren (1999) were not the only authors who found that students tend to assign grades for their peers that were clustered around the mean with very little variation. Goldfinch, Laybourn, MacLeod and Stewart (1999) also found very limited variation in scores provided on one's peers within their teams. Topping (1998) reviewed 25 articles that compared student provided peer evaluations to ratings provided by their instructors. He found that many studies reported that their peer evaluations clustered around the median. Therefore these researchers were in agreement that student evaluations pushed the poorest and highest performers toward the middle.

Students may not accept the concept of other students rating them. Poor group performers may not believe the evaluations they receive from fellow students (Topping, 1998). The conclusion here was that any time an instructor requires a peer evaluation, the instructor must stay involved in the process. Initial instructions followed by continuous monitoring of the situation is required to uncover and fix problems as soon as possible (Topping, 1998).

The grade a student normally receives on their individual projects may differ significantly from the grades they receive on a group project. There are several forms of traditional grading including tests and group projects. A relatively new means of grade determination used in some business schools is the assessment center. The idea behind assessment centers is the use of different exercises that are a reflection of what a manager might experience during his or her normal business day. A major flaw in using traditional forms of assigning grades is that they only determine the declarative and the knowledge compilation earliest stages of learning. Where the assessment center captures all stages of learning including the elusive proceduralization stage that allows application (Bartels, Bommer, & Rubin, 2000). The researchers determined whether or not there was a relationship between 347 undergraduate students' GPAs and how they did on an assessment center. The researchers make it clear that an individual's GPA is impacted by many different things such as motivation and interest instead of simply a matter of intelligence. However, they found that scores on the assessment center correlated with most other scores provided during the class including GPA, tests, and discussion. The only scores that did not correlate with the assessment center score were the scores they received for the group projects. The reason provided for a lack of correlation on the group project was

the fact that the group grade was not done on an individual basis where the assessment center grade was always individually graded (Bartels, Bommer, & Rubin, 2000).

Cooperative and Group Tests

Some research studies show that instructors can further improve the quality of student team learning through the use of group tests. Guest and Murphy (2000) state that prior "...research on the nature of memory suggests that traditional written individual examinations may not maximize long-term retention of information and concepts" (p. 350). They studied 90 graduate students in a teaching program. The researchers developed a group verbal final exam that determined mastery of course materials. They collectively designed the group test so that it required student application of key points from throughout the term. The researchers compared the results from the group exams to students who took the test individually. During the group exam, any student could be called on to respond to any question and their response would be the only grade their group received for that particular test question. At the end of the exam, all members of the group provided critiques of the test format. From these critiques, less than ten percent of the students had any negative comments and the negative comments were generally not related to the group nature of the test. Rather, the

criticisms were centered on the additional pressures the students felt in responding for the entire group. The vast majority of students reported that they felt they would retain more of the information from this testing format than they would from other means of learning (Springer, Stanne, & Donovan, 1999).

Specific findings from the Springer, Stanne, and Donovan (1999) study were that students who were in the group test classes reported spending significantly more time preparing for the final exam than students who were in an individual setting. The authors felt this was a result of the interdependence the students felt and not wanting to let down their team members. Further, the authors felt the extra time spent studying aided retention and contributed to the students reporting that they felt they would retain the information more than in other classes. Finally, the researchers indicated that although there were significant benefits to a group test, professors should not use the group exam as a significant means of assigning student grades. Instead, the group test should simply be one of several different measurements.

Hite (1996) conducted a study of 278 undergraduates by separating them into fairly equally divided control and experimental groups. The control group took three course tests and one final examination as individuals. The experimental group accomplished the exact same tests as the individuals accomplished.

The difference the experimental group received was that they also accomplished the same three course tests in a group setting on the day after they took the tests as individuals. Instead of repeating the exams, the control group spent the entire next in-class session reviewing the tests. Therefore, the amount of in-class time spent on the tests was about equal. The experimental group took the final exam on an individual basis and did not have the same test review as the control group. The experimental groups were comprised of three students with one high, one middle, and one low performing individual in each group. Composition of the groups changed after each test based on the scores from the previous test(s). All students took this final on an individual basis. Although the final exam was comprehensive, no test questions were repeated for either the control or experimental groups. Results clearly showed that the experimental group did significantly better on the final exam than the control group. The main conclusion reached in this study was that the students learned more and had better retention from the group nature of the previous tests. In addition to differences in final exam grades, Hite (1996) found that the end of term student critiques from the experimental group was more favorable than from the control group.

Team Learning

"What students learn is greatly influenced by how they learn, and many

students learn best through active, collaborative, small-group work inside and outside the classroom” (Springer, Stanne, & Donovan, 1999, p. 21). Learning within teams is effective because small numbers of students work together and help one another succeed. It is the significant positive influence brought to bear by the students’ peer group in an academic setting that provides the foundation for improved learning and retention. Cooperative learning is more effective than traditional means of teaching, in part, because team learning motivates and actively involves the individual in their own education. In addition, a group learning environment provides the student with the opportunity to see how their peers handle the same situations they are exposed to. Seeing how others problem solve gives the student the chance to adjust their own techniques and modify their views when presented with conflicting view points (Hite, 1996).

Springer, Stanne, and Donovan (1999) found that students derived several significant benefits when they were taught in a group setting. Not the least of these findings was that undergraduate science, mathematics, engineering, and technology students had a 22 percent higher rate of continuing in their major if they were taught in a group. Such an increase is important because in the science, math, engineering, and technology fields, there is an exceptionally high rate of students who change majors.

Springer, Stanne, and Donovan (1999) also found that undergraduates who were in a small group setting did statistically significantly better than individuals who were not in a group learning environment. Further, these researchers found that students from the group learning classes reported more favorable feelings toward the subject matter they were exposed to. In part, the reasoning behind their findings was that the students may place more value on the group succeeding, may support one another more, and would learn more from one another in a group setting than individually (Springer, Stanne, & Donovan, 1999).

There were no differences in the successes of male and female students within group learning courses. Further, African American and Latino students did better in undergraduate science, mathematics, engineering, and technology classes than those same minorities who were not in a group learning environment (Springer, Stanne, & Donovan, 1999). Their study also found no difference between cooperative and collaborative group teaching styles. The authors defined the cooperative approach as being more instructor structured than the collaborative approach, which relies more on the groups to determine how they will accomplish the assigned tasks. The important thing was not the type of teaching style that was used for the group but rather that there was a group (Springer, Stanne, & Donovan, 1999).

When students are assigned to teams without prior team training, they may not learn as much as if they were first taught about group learning and then allowed to work within a team. Further, many instructors are not totally sure how to properly assess the accomplishments of a team or the individuals that make up the team. In a 1999 study, Goldfinch, Laybourn, MacLeod and Stewart addressed all of these problems. The authors first taught the students what teams were and what makes them function effectively. Next, they brought in several local employers and trained them in what they should be looking for during future periods of student assessment from an individual and team perspective. Lastly, the authors used the employers to provide feedback to their students while the students were still at an early stage in their group projects. The perspective of the employers was found to be important to the students because they had an air of realism. In addition, they had gained the immediate respect of the students because they were already assessing employees who were working in teams. The major findings from this study were that the students seemed to more readily accept the team concept after they received training and feedback from the practitioners. Further, when the same students were seen in later classes, they were actually using the principles they learned about working within groups in the earlier classes (Goldfinch, Laybourn, MacLeod & Stewart, 1999).

The majority of students reported improved learning when they worked within groups. In a study where 140 undergraduate computer science students were equally divided between group and individual learning situations, those assigned to the groups reported that they felt they had learned more. Although the students in the study reported that they had learned more, their grade distribution showed there was no difference between group and individual learning (Benbunan-Fich, 1999).

References

- Bartels, L. K., Bommer, W. H., & Rubin, R. S. (2000). Student performance: Assessment centers versus traditional classroom evaluation. *Journal of Education for Business*, 75, 198-201.
- Benbunan-Fich, R. (1999). Assessing learning of asynchronous computer-mediated communication in the classroom. *The Journal of Computer Information Systems*, 39, 82-87.
- Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment and Evaluation in Higher Education*, 24, 301-314.
- Dyrud, M. A. (2001). Group projects and peer review. *Business Communication Quarterly*, 64, 106-112.
- Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment and Evaluation in Higher Education*, 24, 365-377.
- Ghorpade, J., & Lackritz, J. R. (2001). Peer evaluation in the classroom: A check for sex and race/ethnicity effects. *Journal of Education for Business*, 76, 274-281.
- Goldfinch, J., Laybourn, P., MacLeod, L., & Stewart, S. (1999). Improving groupworking skills in undergraduates through employer involvement. *Assessment and Evaluation in Higher Education*, 24, 41-51.
- Kelmar, J. H. (1993). Peer assessment in Graduate Management Education. *The International Journal of Education Management*, 7, 4-6.
- Kerr, P. M., Park, K. H., & Domazlicky, B. R. (1995). Peer grading of essays in a principles of microeconomics course. *Journal of Education for Business*, 70, 357-363.
- Hite, P. A. (1996). An experimental study of the effectiveness of group exams in an individual income tax class. *Issues in Accounting Education*, 11, 61-59.
- Maranto, R., & Gresham, A. (1998). Using "World Series shares" to fight free riding in group projects. *PS, Political Science & Politics*, 31, 789-791.
- Persons, O. S. (1998). Factors influencing students' peer evaluation in cooperative learning. *Journal of Education for Business*, 73, 225-229.
- Orpen, C. (1994). Perceived similarity: Its effect on the accuracy of peer evaluations among university students. *The International Journal of Education Management*, 8, 4-6.
- Guest, K. E., & Murphy, D. S. (2000). In support of memory retention: A cooperative oral final exam. *Education*, 121, 350-354.

- Lejk, M., Wyvill, M., & Farrow, S. (1999). Group assessment in systems analysis and design: A comparison of the performance of streamed and mixed-ability groups. *Assessment and Evaluation in Higher Education*, 24, 5-14.
- Sherrard, W. R., Raafat, F., & Weaver, R. R. An empirical study of peer bias in evaluations: Students rating students. *Journal of Education for Business*, 70, 43-47.
- Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69, 21-51.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249-276.

Writing as a Tool for Teaching and Evaluating Student Performance

In University Level Course Work

Dr. Nina Haydel

Abstract

This paper addresses the need for writing as a component of every discipline, and the value writing has for both the students and the instructors. The paper presents an historical background of composition in higher education and the rationale behind using writing as one of the methods of educating the students at ERAU. The need for writing and its value as an educational tool in even the most technical of disciplines are explored. The paper attempts to answer three questions that plague instructors when considering the addition of a writing component to a content course:

1. Why do I need to add writing, since my course is technical, theoretical, or hands-on?
2. With all the work I am doing, when can I find time to deal with a written assignment?
3. How can I possibly teach, evaluate, and grade something else, when it seems irrelevant to what I am teaching?

Included are implementation of strategies, management of time, styles of writing, suggested assignments, and performance evaluation/grading/assessment advice for both formal and informal, graded and un-graded written assignments, along with sample writing criteria.

We do not think and then write, at least not without putting an unnecessary handicap on ourselves. We find out what we think when we write, and in the process put thinking to work-and increase its possibilities.

-Frank Smith (1982)

Introduction

Embry Riddle Aeronautical University professors are dedicated to the goals set forth by the university: "...to provide a comprehensive education that prepares graduates for productive careers and responsible citizenship to support the needs of aviation, aerospace engineering, and related fields" (ERAU Graduate Catalog, 1998-1999 p. 7). Degree programs focus on varied aspects of the aviation industry, generally with a technological emphasis in both undergraduate and graduate programs. Along with the technical courses, the core programs include several composition courses. Those of us who teach writing courses are fighting a battle to facilitate our students' communicating through the written word. We seem to be losing that battle. We cannot do it alone, so we need to look to other members of the faculty to provide the missing components of this struggle to present Standard Written English as a means of communication and to provide another avenue for student learning, while using writing in different academic settings.

Historical Background

Written communication, today, is one of the bedrocks of our society. Cave walls were adorned with figures representing meaning; parchment served to record what we now consider valuable information; paper continues to provide a trail of important ideas; the computer has usurped the spotlight as the major tool for recording the written word. Writing has not always been a part of our major means of communication. Before the 1870s, oratory was a primary focus in education. Schools and universities concentrated on training students in rhetoric and public speaking, for recitation was the standard. Writing activities centered solely on classroom note taking. If someone spoke "correct English," he or she was regarded as having the capacity to present a high standard of written literacy. Unfortunately, many students who were judged on their writing skills did not measure up to the standards expected of their professions or social class (Russell, 1991).

When university education became more practical and more easily obtainable, universities engaged in research and began to hold campus-wide oral and written examinations graded by faculty, collectively, or by outside assessors. Writing grew to be regarded as a skill independent of content learning. Gradually, as universities developed and grew in complexity, composition courses were accepted, but there were no provisions for college-wide

writing outside those composition courses.

Junior faculty and teaching assistants assumed responsibility for the writing courses, and no other faculty participated in the perpetuation of the skills students acquired in the dedicated composition classes. Styles of writing included class notes, laboratory reports, research papers – all which placed significant value on discovery and factual evidence, thus devaluing the “the written form of intellectual discourse”(Russell, 1991). Students and faculty focused on academic writing that presented content and ignored the context of the writing with regard to audience and purpose, thus not allowing for differences in the ways disciplines presented the content.

During the Progressive Era, education developed into a more utilitarian institution and viewed the goal of writing instruction as a means to provide skills needed to follow instructions and develop the relationship between the writer, reader, subject, and point of view. Following World War II, knowledge exploded and multiplied, accelerating the pace of change (Russell, 1991). With the advent of the Information Age, writing became viewed as less important. Short emails and instant messages seem to have usurped the letter writing and instruction delivery of the recent past. That is where we are today. Academia seems to have no shared vocabulary and no forums for dealing with discipline-specific writing instruction.

The Problem

We, as faculty of Embry-Riddle Aeronautical University (ERAU), can easily find ourselves inundated with student papers (often unworthy of our reading time), particularly if our courses require multiple writing assignments; therefore, many instructors shy away from making written assignments in classes of a technical nature. Therein lies the problem: Without adequate practice and refreshing of writing skills, students in university level courses are inadequately prepared to deal with the written word, when it is a required means of communication.

As a technical institution, ERAU has the reputation for preparing its students to successfully enter the employment realm in their respective, specialized aeronautical areas. Students are required to major and minor in various fields, as well as complete the core requirements for a liberal education. Numerous courses are available, some mandatory, some elective, as part of the Humanities Department, including several writing courses. During these classes, students focus on a variety of writing skills, which, unfortunately do not always transfer to other areas of their educational experience.

The Need for Writing in the Disciplines

A major need exists to help alleviate that situation. Writing teachers cannot “go it alone.” Every instructor, regardless of subject content needs to become a writing teacher, as well. This can easily be

accomplished if all instructors build into their assignment and grading plans the opportunity for all students to continue writing. It can stimulate dialogue between the student and the instructor, particularly in a large class. The importance of this addition to all curricula cannot be overestimated. To hold students accountable for their written assignments will enable the instructors to enhance their own educational productivity and improve written assignments throughout the university.

Three major concerns often surface with regard to one's adding a writing component to any course:

1. Why do I need to add writing, since my course is technical and hands-on?
2. With all the work I am doing, when can I find time to deal with written assignments?
3. How can I possibly teach, evaluate, and grade something else, when it seems irrelevant to what I am teaching?

Question 1: Why do I need to add writing, since my course is technical, theoretical, or hands-on?

Rationale

Although writing serves many cognitive purposes, it primarily serves as a mode of learning. According to Janet Emig (1997), writing in all content areas trains students to think about a topic, internalize the topic and reach their own conclusions, thus causing

them to analyze, compare, and synthesize relevant facts and material. Students often view knowledge as the acquisition of correct facts and information, when, in reality, true knowledge is dynamic and requires a "complex combination of potential solutions arrived at through critical thinking" (Schmidt). There is a significant link between critical thinking and writing. Students can develop intellectual and cognitive abilities when thinking about important problems that form the center-point of a writing assignment. Once they begin writing, they engage in expanding, clarifying, formulating and deepening their thinking. As quoted in Schmidt, J.C. Bean states, "When we make students struggle with their writing, we are making them struggle with thought itself." Writing assists students in learning content more effectively by communicating the knowledge, regardless of discipline.

Writing is a tool, not only for learning, but also for discovery. Unfortunately, writing has been used as an end product or a demonstration of something that has been learned. This type of writing is considered the product-centered model and the prove/approve model (Griffin, 1983). Writing, according to the Writing Across the Curriculum (WAC) model, is both a process and a product, which integrates the use of eye-hand coordination and involves both hemispheres of the brain. This process enables the student to formulate and connect ideas that he or she can transmit to

others, but can transmit to the learner through the use of higher order cognitive skills necessary for critical thinking (Emig, 1977). Students better focus and understand subject matter; therefore, those who write in the content area often produce better written products with the practice writing affords. Although the writing certainly may not be the primary objective, it is definitely a result (Miller, 1991). According to a three-year study by Judith Langer and Arthur Applebee, funded by the National Institute of Education, "...the more that content is manipulated through writing, the more it is likely to be remembered and understood" (Langer, J. 1992, p. 130).

Regardless of discipline, writing has a place in each curriculum. WAC is a student-centered pedagogy, not a single movement or a trend, in which all disciplines value writing, rather than considering it an additional burden. Writing is a significant part of learning and serves as a binding force between the creation and acquisition of knowledge (Russell, 1991). Hence, along with teaching content, the instructor becomes a writing coach who can help students organize, synthesize, evaluate facts, and communicate in such a way as to demonstrate learning in that particular discipline, within the conventions of the written word. Since higher education institutions do not just provide instruction, but produce learning, passive learning becomes active discovery, exploration, and invention (Schmidt). Writing often has some

of its best applications in disciplines such as physics and mathematics, where students need to learn to think through and evaluate problems, thus becoming one of the most powerful tools students can use to learn how to organize their thoughts, decide what is important, and enable them to know what they really know (Cooke, 1991; Heyman, n.d.).

Question 2: With all the work I am doing, when can I find time to deal with a written assignment?

Time Constraints

You, the instructor, are probably wondering about the paper load you think will be added to your already full schedules. You are probably thinking that you are not English teachers and do not want to do our job. Yes, student writing can be very time consuming, but you need not allow it to be when teaching your classes. Yes, it will take away from your limited time in class, but only in a very minor way. Adding writing assignments in content-driven courses is difficult, especially when students resist anything they see as being unrelated to course content. You must wonder if it is worth the effort? Rest assured, it is!

In just a few minutes, with appropriate writing activities, you can gain valuable insights about problems your students may be having with content, about what they do or do not understand, about how they feel regarding certain issues. This can help you make decisions about how to proceed with

your content or what you may need to review. You will not be teaching writing, but you will be using it to help your students learn the content. You do not need to know how to correct student errors, as you will be responding to content. If you do respond to errors, you need not necessarily correct them; your response to errors represents the value you place on writing and the quality you expect from your students. You need to be able to understand what they are trying to tell you. If you expect more, you will receive more!

When students know you value the written word, they will extend themselves to make their ideas clearer and more comprehensive, not forcing you to do the work of perceiving what they are trying to say. Require students to write multiple drafts. The writing process can be accomplished outside of the classroom. Request rough drafts of all written work; that will cut down on plagiarism. You need not read their rough drafts, but a minute to glance holistically over a paper will give you enough insight regarding the students' accomplishments. Require drafts submitted at an appropriate time before the paper or assignment is due. You will be aiding the student in time management, as well as receiving your assignments when due.

There are many ways to insert writing into your schedule without taking much of your time and energy. First, ask yourself what you want the students to learn from your assignment. Your learning objectives,

frequently articulated on the basis of *Taxonomy of Educational Objectives: Cognitive Domain* (Bloom et al, 1956), can be used to establish how you wish to assess what you want your students to accomplish.

Example: Lower order cognitive thinking- knowledge of the topic
Higher order cognitive thinking- solution to a problem, evaluation of class discussion

You will find that students learn key concepts and understand material more fully when they write about the content. They think about the topic and internalize it in order to reach their own conclusions, which they can express more clearly and more thoroughly. Often, after presenting a new concept and asking for questions from the class, you may be facing blank stares. Consider asking the students to write what they understand you to have taught. You may discover the student who is shy about speaking up in class will express his or her confusion in a more private way, more willing to admit a lack of understanding through writing.

An action research study in 1989, by Miller and England, found that allowing students to write five minutes in four out of five instructional sessions did not interfere with the curriculum content over an academic semester. In fact, the writing may have helped the instructors cover the material more effectively (Miller, 1991).

Class time is precious, and every instructor

is loathe to introduce something that seems to usurp teaching time, but writing may help to make your class time more valuable for both you and the students. Writing can add to the richness of your course; you cannot afford **not** to give your students the opportunity to write.

Question 3: How can I possibly teach, implement, evaluate, and grade something that may seem irrelevant to what I am teaching?

Teaching

Regardless of the discipline in which you teach or the course objectives you seek to achieve, writing can easily play an important part of the content of your course. You need not teach writing, although a brief (10 minute) review of some basics (such as APA format) would make life easier for you if you do assign papers. A quick mini-lesson can alert your students to "fragmentitis," a disease that stunts the growth of sentences and ideas. Many can **hear** their problems if you encourage them to read aloud what they have written. You will be using writing to help your students listen to what they know and learn your subject matter, while you can assess any difficulties. You will be giving your students an opportunity to ask for help and become more aware of what they do not know. Promote correct spelling, punctuation, and grammar from all students.

Implementation

Implementation fully depends upon your content and your course objectives. Some general ideas include journals, free-writes,

short reaction papers, generic or focused summaries, lab journals, response papers, learning logs, letters, pre-test warm-ups, analyses of events and processes, project notebooks, discussion starters, using cases, annotations, etc. You can turn a lesson into an inquiry session by having students write ideas that intrigued them or confused them in a chapter or class discussion. Asking a student to prepare a set of instructions or an explanation of a task, recently learned, regardless of the technical nature, will help you assess his or her level of understanding. Students can write procedures and product or equipment descriptions, definitions of concepts, explanations of technical instruments, or short position papers (Haydel 2001). In a class that involves computations, if students write clearly about the concepts, they probably understand them. Writing about how students approach computations allows them to think more sharply, thus revealing a comprehensible understanding of a process (Emig, 1977; Miller, 1991).

Impromptu Writing

Impromptu writing prompts (starter phrases, ideas, concepts) constructed by the professor, regardless of the subject matter, solicit written responses to specific problems or questions as a valuable teaching tool. You may want to solicit clear explanations of a concept, skill, or generalization.

1. Set a time limit for any in-class writing assignment. Stick to it (three minutes, five minutes).
2. Prepare the impromptu topic in advance.
3. Be flexible. Sometimes a “teachable moment” relates to something unplanned. Seize the moment and use that idea for an impromptu topic.
4. Since students seem to write better if they know they are writing for a particular audience. Establish the audience in advance (i.e. peer, supervisor, administrator).
5. Have the writing take place at the beginning of the class to serve as a transition into your subject.
6. Have students write at the end of class to assess their understanding or impressions.
7. Write with the students. Write a response to the prompt from the professor’s point of view (if appropriate).
8. Discuss the purpose of the writing experience (perhaps to serve as a more effective learning tool), so the students will not expect anything other than intrinsic rewards.
9. Do not penalize students when they do not write. Writer’s block hits all of us at times.
10. Initial writings may be brief and meaningless at first. Be patient and give students an opportunity to

develop the skill and reap the benefits of writing in your discipline. (adapted from Miller, 1991, pp. 516-521).

“Think on Paper” Strategy

1. Explain the purpose of the writing before assigning the paper
2. Focus on depth of thinking and quality of ideas.
3. Spend as little as five minutes on the actual writing.
4. Use this to precipitate class discussion.

Micro-themes

Micro-themes serve many purposes and have numerous advantages. They require significant thinking time, but little text. This snapshot into students’ thinking processes maximizes student learning and minimizes instructor-grading time. Four styles of micro-themes are highly productive: summary writing, thesis/support, data description, and quandary posing.

1. This mini-essay can range from a few sentences to a few pages. It may extend from serving as a learning tool as a way to summarize lecture highlights, to accessing student questions, or generating class discussion.
2. This activity enables students to become active class participants.
3. Microthemes can be so short they can be typed on one 5” X 8” note card. An example is an article summary in fewer than 200 words.

4. Use this format as preliminary, exploratory writings.
5. Microthemes usually are not graded. They can also become the basis for more formal, thesis-governed academic writings that are graded (Schmidt).

When requiring a formal technical proposal, laboratory report, discipline-based academic report, several suggestions will help you avoid student difficulties.

Suggestions for Presenting Formal

Reports:

1. Be clear about your expectations. Present all instructions in writing. Advise the student of the purpose of the written assignment.
2. Establish a specific audience for the paper to direct the student to appropriate terminology and format).
3. Explain the kinds of sources you expect (differences between primary and secondary sources, annotated bibliographies).
4. Use the idea of the writing process that includes composing (drafting and outlining ideas, establishing a working hypothesis) drafting, revising, editing.
5. Explain the conventions of an academic product and what you expect (parenthetical references, bibliography, APA style- it has recently changes, so you will have to update yourself).

6. Set up your criteria before the assignment and alert the students to your grading policy. Make students aware of the value you place on their writing, and generate a concern for correct Standard Written English. Build the writing clarity and correctness into your grading policy, or students will not value their writing skill.
7. Provide samples of excellent and not-so-excellent papers, so students can see what to do and what not to do.
8. If possible, allow students to submit drafts for perhaps a peer evaluation session before the final paper is due and encourage students to revise their writing.

You will probably find it preferable to assign several short pieces of writing rather than one long one. If you ever have time (doubtful), try one of your own writing assignments. You might discover some surprises about your expectations. Consider the objectives of your course and tailor the assignments accordingly. Sequence your assignments, beginning with something students know how to do; then build others based on previous assignments (Walvoord et al, 9, 14). Allow students to do un-graded exploratory writing to help them develop ideas that can be used to further discussion. Include in an exam a previous "quick write" that had been done in class, to show the value of thinking on paper.

Some Validating Studies:

Several studies validate the success of writing across different disciplines throughout the United States.

1. George Mason University in Virginia reports the following: Business statistics students who keep journals averaged 10 percent better in their final exams than those who did not (Miller 1991).
2. Montana State University reports: Eighty-eight percent of students surveyed thought writing had helped them understand physics (Miller, 1991).
3. Rutgers University in New Jersey reports: Math students who verbalized their difficulties in writing were often able to understand problems they could not solve before. (cited in Miller, 1991, Madigan's "Writing as a Means, Not as an End")
4. A study by Stout, Wygal, and Hoff (1990) of Rider University, revealed many professors enhanced the writing skills of their financial accounting students by emphasizing the importance of writing correctly and required them to maintain journals and complete a small-group formal writing assignment.
5. Hall and Tiggeman (1990) required their students in an introductory finance class to write short, frequent, informal assignments

which improved students' understanding of finance concepts and enabled instructors to better assess their students' comprehension (Miller 1991).

6. Toyota Motor Manufacturing, USA, expects all its employees to write one-page reports using graphics (Carnes, 2001).
7. A survey of 52 engineering firms revealed that "writing proficiency is a major factor in deciding the promotional potential of an engineer" (Sully, 1995).

The essence of writing assignments, regardless of length, depth, and formality provides our students with an opportunity to learn how to solve problems, examine ideas and support them with evidence, incorporate and synthesize information, and transmit clearly to others. Language proficiency fosters economic and social opportunities.

Performance Evaluation/Grading

A common complaint by professors of every discipline in every educational institution focuses on grading overload. Now, with the added responsibility of instituting writing in your course, you are probably shuddering about adding a new grading component. You can establish something as simple as the following: If you cannot understand what the writer is saying, lower the grade by the number of points you deem appropriate. Always establish your grading policy when you make your assignment to include Standard Written

English. Remember, you need not grade every writing assignment. Decide in advance as to the purpose of the writing and alert your students to the value you place on each aspect of the product. Put every item and its value in writing.

Evaluating a Formal Paper for a Grade:

1. You can easily skim to understand context; this avoids your asking questions, which may be clarified on later pages.
2. Don't edit the student's work. Select one paragraph or one page and identify spelling errors and grammar problems. Write a note telling the student the rest needs cleaning up and editing.
3. Look for the important points that you have established in advance. Be an active reader by underlining items you feel are appropriate or placing checks in the margin adjacent to the items.
4. Point out strengths in the content as well as weaknesses or gaps.
5. If you cannot understand an idea, ask questions that reveal why you are confused. Often, the writing or misuse of words and punctuation causes the confusion. Alert the writer to what you see (Cooke, 1991; Haydel, 1987).
6. You can establish a step-by-step method where you present students with dates for submission of segments of the assignment. In that

way, you can reduce the burden of reading a lengthy work. Suggested activity:

- a. Students may submit an hypothesis/study question or rationale, depending upon the assignment. This may take the form of the paper's introduction and be graded.
 - b. Students turn in outline of ideas and annotated bibliography.
 - c. Students may submit final draft of body for grade.
 - d. Students present final draft of conclusion for grade.
7. Develop an evaluation checklist that identifies what you are looking for in the assignment. Be sure to include readability in the list and assign points for lack of clarity and understandability.
 8. Base the grade on previously announced criteria. Weigh each criterion appropriately.
 9. Assign a grade that reflects the document's usefulness in the real world or the world you have established for the assignment. Has the document achieved its objectives?
 10. Provide feedback concerning the content. Look for positives as well as negatives.

Simplified Evaluations of Writing

Assignments

You can conquer the grading overload by following several suggestions that will simplify the evaluation process:

1. Keep assignments brief. You may find that one or two paragraphs may be as effective as one or two pages.
2. Do not be an editor. Students will not learn if the work is done for them.
3. Build in a self-evaluation component to an assignment. Allow the student to assess his or her own work.
4. Don't be concerned if you cannot identify the writing problem. If it does not sound right, it probably is not. Just note it as a problematic section of the assignment. Lack of clarity blazoned upon a paper sends a message to the student.
5. Develop a peer evaluation session. Present students with pre-designed forms to be turned in with the assignment. Use pairs to evaluate one another's work with respect to your instructions and criteria.

Un-graded Writing

Writing need not be graded. Do not respond or grade informal writing, but comment on the content. Drafts need not be graded, but they should be assessed according to content requirements. You can use writing as part of your evaluation process, yet not as a separate grade. Un-graded writing can serve as an informal

assessment of student understanding, which will give you a clearer picture of areas that need remediation. There will be occasions where you do not need to write comments to a student; you can say something to the entire group or to a specific student, allowing the students to know that you read their writing. Your form of response is secondary to the concept that you are providing feedback.

Treat much of the student in-class writing as exploratory work, thinking on paper. Resist the attempt to judge too much. You can select several anonymous papers and read them aloud to the class, as examples of your expectations. Comment on common patterns and interesting ideas; pose questions not explained by the writing. Use this as an informal way to get information from students about the ideas that need reviewing or expanding.

Suggested Evaluative Comments for Student Writing Assignments

Two types of comments are most relevant to students' written assignments, atomistic comments and holistic comments. Atomistic refers to grammatical, punctuation, diction, mechanical, vocabulary, and spelling errors. Holistic refers to organizational problems, clarity of ideas, confusion of concepts, illogical conclusions, insufficiently supported assumptions, and lack of evidence (Haydel, 1987).

If atomistic errors proliferate, select one page and identify the most obvious. You need not be a grammarian to see these

common mistakes. Focus on the holistic problems that cloud the meaning of the content. Frequently, a confused paper represents a confused mind.

Instructor comments are most instructive when they are presented in the following way:

1. Use vocabulary that the student will easily understand. Avoid technical references in your comments.
2. Provide complete comments, rather than cryptic phrases.
3. Always provide comments in the graded, final draft.
4. Try to find something positive to say.
5. Give specific advice concerning conceptual problems.
6. Mention editing problems, but focus on ideas.
7. Encourage students to use peer revision. Other eyes see things the writer misses.
8. View responding as a means of providing feedback instead of as a means of justifying grades.
9. If possible, make comments during the writing process, rather than after the final draft. Comments redirect and improve thinking, which leads to improved writing. Grade only the final product (Schmidt).

Sample of Grading Criteria

Rubrics are often created to focus on the expectations of the assignment. Criteria should match the assignment objectives.

Grade of A: Consistent, clear, effective communication of ideas.

Communicates purpose to audience.

Clear focus, adequate development, coherent thoughts.

Appropriate organization for topic.

Provides required material as determined by assignment.

Grade of B: Strong, but contains occasional deficiencies in one area, with minor problems in several

areas.

Distractions caused by errors in writing or style.

Grade of C: Does the job adequately.

Deficient in more than one area.

Maintains its purpose.

Grade of D: Serious problems in focus.

Lack of development.

Unable to convey message.

Purpose is undermined by errors in writing or style ("A fuller Definition).

Do not implement writing in all classes at the same time. Use it as a tool when it is relevant to what you are teaching. Stagger your workload. Holistic reading will allow you to read as many as 25-30 short responses in 5-10 minutes. Look for specific responses;

some will write more than others; some will write almost nothing. Consider writing quickie, non-judgmental responses when appropriate.

Conclusion

The ability to write is fundamental to success. Writing shapes thinking, facilitates learning, and empowers creativity. We at ERAU must successfully prepare our students for the changing workplace of the 21st century. Students need to learn how to solve complex multi-disciplinary problems through effective communication, which goes beyond merely expressing oneself orally. Education must extend beyond the acquisition of knowledge and facts to include critical thinking and clear, logical communication of information.

Serious problems occur when a professor cannot distinguish between students who understand the content but have trouble communicating it and those who really do not understand. Poorly written work is difficult to grade, since it obscures the content. Students should be held responsible for their ability to communicate their ideas, right or wrong. Perhaps students do not perceive writing as critical to their success and are detail-sloppy; perhaps students do not spend enough time on their assignments, waiting to the last minute without time to revise. We need to refocus them so they will understand the need for better language skills and understand that content is not the only component of success. When content instructors pass on

the responsibility of writing to others, they are, in effect, signaling that only the content is important, but the process is not.

It is our job, as ERAU instructors, regardless of our disciplines, to assist students in their educational process by providing writing as a means of solving typical academic problems and developing students' analytical and communication skills. Our society demands that, regardless of their vocation and technical skills. Writing is an integral part of learning, regardless of subject matter.

References

- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Kralhwohl, D. R. (1956). A taxonomy of educational objectives: Handbook 1, the cognitive domain. New York: David McKay Company.
- Carnes, L. W., Jennings M. S., Vice, J. P., & Wiedmaier, C., (2001). The role of the business educator in writing-across-the-curriculum program. Journal of Education for Business, no. 4, March/April, 216-219.
- Cooke, V. (1991) Writing across the curriculum: A faculty handbook. Victoria: Centre for Curriculum and Professional Development.
- Embry-Riddle Aeronautical University Graduate Catalog (1998-1999), 7.
- Emig, J. (1997, May). Writing as a mode of learning. College Composition and Communication, 28, 122-128.
- A fuller definition of writing to learn. (n.d.) Colorado State University. Retrieved April 19, 2002 from World Wide Web: <http://a.w.colostate.edu/reference/wac-faq/pop4a.htm>
- Griffin, C, & Williams, E. (1982) Teaching writing in all disciplines. San Francisco: Jossey-Bass.
- Haydel, N. (2001). ERAU syllabus for COM 221, technical writing.
- Haydel, N. (1987). A study to determine the value of computers in the revision process of written communication (Doctoral dissertation, Temple University, 1987).
- Heyman, G. A. (n.d.) Writing to learn, great ideas for teaching accounting. Retrieved February 16, 2002 from World Wide Web: http://www.swcollege.com/vrcomm/gita/gitta22-5_main.html
- Langer, J. (1992), Speaking of knowing: Conceptions of understanding in academic disciplines. Writing, Teaching, and Learning in the Disciplines, Herrington A. & Moran, C. (Eds.) NY: MLA.
- Miller, L. D. (1991). Writing to learn mathematics. Mathematics Teacher 84, 7, 516-521.
- Russell, D. R. (1991). Writing in the academic disciplines 1870-1900: A curricular history. Carbondale: Southern Illinois University Press.
- Schmidt, S. J., (n.d.) Using writing to develop critical thinking skills. University of Illinois. Retrieved March 19, 2002 from World Wide Web: <http://www.english.uiuc.edu/cws/wac/schmidt.htm>
- Stout, D. Wygal, D. & Hoff, K. (1990). Writing across the disciplines: Applications to the accounting classroom. Bulletin of the Association for Business Communication, LIII (4), 10-16.
- Sully, B. (1995). Malaspina University-college's writing-across-the-curriculum project, www.mala.bc.ca/www/wac/proj/.htm, retrieved 2/16/2002.
- Walvoord, B.E. & McCarthy, L.P. Thinking and writing in college: A naturalistic study of students in four disciplines. Urbana, Illinois: National Council of Teachers of English, ERIC ED 324591.

**WRITING IS A VERB, NOT A NOUN:
USING EVALUATION TO ENHANCE STUDENTS' WRITING SKILLS**

by
Patricia Valley

ABSTRACT

Listen long enough to faculty members at almost any university in the United States, and one will hear of the dismal writing skills exhibited by all too many students. Clearly, students who write poorly are ill-equipped for today's demands in industry and business, and their lack of writing ability may contribute to difficulty in reflecting on course content and in critical thinking. We know that writing often will not improve within the time confines of a single academic term; lower functioning students in particular tend to develop their writing skill over time. With so much at stake, it is vital that students improve this important skill in every course they take.

What should instructors do to help students improve their writing? What works? This paper presents strategies for the evaluation of student writing in ways that facilitate improvement. Principles of writing evaluation are identified, and insights gained from developmental writing classrooms and writing across the curriculum efforts are discussed. A presentation of informal assessment techniques such as self-evaluation, peer evaluation, and writing for revision describes a number of strategies for helping students to view writing as a process, rather than as a finished product that cannot be improved. Formal assessment of writing topics such as grading, ranking, analytic methods of evaluation, and holistic methods of evaluation, along with rubrics and descriptions of evaluative scales, provide useful tools for assessing written work in any content area. A brief outline of the affective domain and potential measures for student writers is included as well. The myriad of ideas presented here serves as a starting point for those who want to help their students be successful communicators.

**WRITING IS A VERB, NOT A NOUN:
USING EVALUATION TO ENHANCE
STUDENTS' WRITING SKILLS**

Introduction

Evaluating students' written work in ways that encourage improvement is one of the more difficult tasks that instructors face. The writing-across-the-curriculum movement has assumed that the promotion of writing increases students' learning (McGovern & Hogshead, 1990), a concept that has been espoused by many researchers as a means of both helping students learn subject-specific material and of improving their written communication skills (Gruber et al., 1999). Writing in the disciplines should contribute positively to students' overall development: It should help students learn course content, improve their writing, and generate new meaning through reflection. McGovern & Hogshead (1990) affirmed that "writing is a complex intellectual process" and that "writing is a mode of learning as well as communicating" (p. 21).

Some instructors view writing as a skill with parts that may be taught in order to produce a measurable product. According to Zinn (1998), such a view may harm students with lesser writing abilities as the high volume of comments on these students' papers may be unclear and appear arbitrary to the students. Perhaps our definition of writing needs to change: McGovern and Hogshead (1990) changed their views through their involvement in a university project focusing on the definition and assessment of

general education at Virginia Commonwealth University.

Our thinking about writing was fairly basic. [We formerly believed that] Writing is a skill that is learned through opportunities to practice and by receiving feedback from an instructor. We were thinking about writing as a noun, that is, as a text produced by the student and corrected by the instructor. Our emphasis was on traditional assignments and giving feedback to our students about their prose. We began to construe writing as a verb. Writing is an action, a process of thinking and learning, which is inextricably tied to our students' cognitive development in our particular courses and in their college careers in general (p.5).

This paper assumes the value of incorporating writing into the curriculum across the disciplines and the use of writing assessment techniques as a means of improving students' written communication skills; it presents assessment methods that practitioners have developed that work.

The Nature of Writing Assessment

Evaluating students' writing can be very time-consuming and is often perceived by instructors as an arduous and sometimes fruitless task. Additionally, the nature of the assessment process itself is fraught with concerns about consistency of evaluation and appropriateness of the measures. Instructors must understand the principles of writing assessment and employ them appropriately to

achieve the highest potential improvement in students' writing.

Definitions

In this paper, the terms "assessment" and "evaluation" will be used as follows:

Assessment: A communication intended to shape students' performance (as opposed to judgment).

Evaluation: Determining worth by outlining strengths and weaknesses. Forms include both formative measures (without judgment) and summative measures (making a judgment, as in assigning a grade). Thus a formative evaluation is very similar if not the same as an assessment.

Principles of Writing Evaluation

The evaluation of writing must be resolutely tied to the goals of the course. What the instructor values is what should be evaluated. The learning outcomes and evaluative criteria should be clear and should be communicated to the students. The instructor should develop his or her own response for written assignments or essay questions and use the essence of the response as a model for rating students' work. This is a good way to ensure content congruence and clarity of purpose.

A joint task force comprised of representatives from the International Reading Association and the National Classroom Teachers of English (1994) established the principles of writing assessment most likely to encourage improvement in student writing.

These principles included the following:

- 1) Evaluation measures and instructor comments should encourage

students to reflect upon their writing in constructive ways that lead to revision.

- 2) Instructor comments should emphasize what students can do, not what they have failed to do. Do not mark every error.
- 3) Hold the standard for students' writing high to yield high quality results.

What We Can Learn from Developmental Writing Classrooms

While on the surface it may seem irrelevant to study the techniques that work in the instruction of developmental writing students to glean helpful practices for the university-level classroom, the same techniques that are endorsed at the middle and high school levels are practiced in post-secondary developmental classrooms and also in universities. The common goal is to align evaluation with course goals and adopt best practices in enhancing students' writing.

According to Cleland (1995), an associate professor in the Department of English and Philosophy at Purdue University, a large contributing factor in poor student writing is a lack of distance. Some students bring a high competency in oral communication, but little sense of how to establish the rhetoric of distance common to academic writing. These students communicate as though the audience were "right there," and convey their thoughts in writing as they would speak them. This causes the writer to leave out contextual details and elaboration of meaning. Such writing tends to lack coherence, context, and discussion. The

first key to helping students learn the difference between oral communication and written communication with its need for elaborated forms and syntax is to immerse them in writing.

Developing writers need multiple opportunities for revision. They must learn to reflect upon their own writing and to improve the quality of what they have produced. This principle holds true not only for writers in need of remediation but also for all writers. Students at all levels must recognize writing as a process, a state of becoming, as opposed to a noun, a fixed state. An important goal of informal writing assessment is to establish the writing and rewriting process as the norm for all writers.

This instructor has found that writing for revision does help students to produce better papers. What is not clear is whether or not students are able to apply the concepts learned to other writing projects. While intuitively instructors would assert that students gain skill which they apply to future writing tasks, the evidence from the classroom is not always clear. From this instructor's experience in teaching developmental writing, the conclusions are that very low functioning students do not seem to apply the improvement in writing from one assignment to the next. It appears that they are unable to improve their perception of audience distance and mechanical weaknesses in the duration of one term to the point that they are capable of reflecting upon their own writing and revising appropriately without outside feedback.

Students who possess a frame of reference for applying the new writing skill learned appear to improve dramatically and

appear to carry their learning over to new writing projects. It is the opinion of this instructor that students' varying foundations for writing improvement is three-fold: critical thinking ability, established writing ability, and reading background. Writing well is intimately intertwined with critical thinking, as one cannot write well without having something to say (Graham, 1992, as cited in Gruber et al., 1999).

What We Can Learn from Writing-Across-the-Curriculum Efforts

At Northern Arizona University (NAU), the engineering faculty in the College of Engineering and Technology (CET) saw writing as a means of involving students in a community of discourse in the discipline and of developing the needed communication skills demanded in industry (Gruber et al., 1999). The purpose of the program was clear: to develop the communication skill in students that industry requires. At NAU, the CET faculty designed a series of four core engineering courses to narrow the gap between engineering students' educational preparation and the industry requirements for success in the workplace. The program, called Design4Practice, addressed the "discrepancies in communication skills, problem recognition and solving, and ethics and professionalism" (Gruber et al., 1999, p. 423). Cross-disciplinary instructional teams emphasized the communication skills needed for management and the profession, especially technical writing. Engineering faculty were encouraged to "see themselves as writing experts in their disciplines who would be able to work with students on improving their writing

skills" (Gruber et al., 1999, p. 424). After a series of faculty writing workshops, the faculty began sequencing writing assignments in the core engineering courses and using a structured peer review process to help students improve their writing. Faculty scored students' work on a four-point scale (4=high, 1=low), addressing the following global issues:

- 1) Clarity and coherence of structure:
 - a) Was there a clear introductory thought?
 - b) Was there a clear assertion of position?
 - c) Was there a logical argument to support the assertion?
 - d) Was there a concluding thought?
- 2) Correctness and conciseness:
 - a) Were the sentences constructed correctly and concisely?
 - b) Was the paragraph structure logical?
 - c) Did the response address the issue of concern in the prompt? (p. 431)

The papers were ranked and read by two readers. Findings showed that there did not appear to be any statistical growth in the students' writing skills. Students' papers were "largely disorganized and incomplete" (p. 432). Anecdotal findings, however, indicated that there was an improvement in students' desire to improve their writing and that students' writing in their assignments did improve. While the average student score did not increase significantly, "individual students' scores increased significantly. On the other hand, some

students' scores dropped, counteracting the positive increase evident in some student writing" (pp. 433-434). It was found that students took the pre-test more seriously than the post-test, largely because the post-test was conducted immediately following an intensive project.

The CET faculty concluded that students' gains in writing may not be evident within a short period of time and that results of writing-across-the-curriculum efforts should be measured on a programmatic, long-term basis. They further commented that holistic scoring of students' writing was an effective way of evaluating students' work (Gruber et al., 1999). These conclusions are consistent with findings from other WAC efforts.

Informal Assessment

Evaluation does not necessarily mean assigning a grade. Not all students' writing should receive a summative evaluation if students are to master the process of writing for revision. Informal procedures are often more productive, establishing a workshop environment rather than a competitive one where grades instead of learning is the goal. Informal writing assessment should encourage students to take risks early in the writing process. In the workshop environment, instructors become facilitators who help students to recognize the problems in their own writing and to correct them (White, 1994; Zinn, 1998).

One of the key concepts of informal assessment, according to Edward White (1994), a leading expert in assessment, is to give meaningful praise when responding to students'

work. White warned that vague compliments, such as "good job" could be detrimental, so the instructor must be specific. He proposed that instructors should mark and comment upon the clearest, most inventive sentence in the student's work. White suggested that instructors pose questions rather than making statements, as questions may inspire students to reflect upon their work.

Self-Assessment

Self-assessment places more of the responsibility for writing improvement upon the student. Rather than serving as a means of reducing instructor workload, as students sometimes perceive this technique, it encourages self-reliance and helps students to ascribe meaning to the long process of writing as they reflect upon their own work. Zinn (1998) suggested the inclusion of the following elements:

- 1) Ask the length of time the students spent writing (from prewriting to the final revision).
- 2) Ask students to comment upon their most meaningful revision, having students outline the strengths and weaknesses of their work.
- 3) Ask students to set at least one goal for improving in the next assignment.

Peer Evaluation

Instructors should realize that peer evaluation takes quite a bit of class time. However, the technique can be very effective, so the benefits of peer evaluation should be carefully considered before rejecting it.

Reasons to consider using peer evaluation of writing assignments:

1. Students enjoy working together and learn group communication skills.
2. Examples of good papers written by peers are more readily accepted than models written by instructors, as students' papers may set a more realistic goal for achievement than the lofty instructor-produced paper.
3. The editing and revision processes are more accepted when done in a group.
4. Students realize that revision is necessary for everyone; thus they lose the misconception that having to revise equals failure (Zinn, 1998).

Several means of achieving peer collaboration for improving writing exist. They include forming groups in which one student reads another student's work aloud and offers an oral reaction to the work. Another technique involves having student groups comment on each other's work in specific areas, adding suggestions for improvement. Another method espoused by Elbow (1973, as cited in Zinn, 1998) uses summary techniques. One student reads his/her paper aloud, and the responding student has fifteen seconds to name the main points of the work, using different words than those used by the writer. Then, the responder or the group attempts to summarize the work in one sentence. Finally, each member of the group chooses one word to summarize the paper. Through this process, the writer will see whether or not the intended meaning of the paper came through. Peer processes should be highly structured for the first few times students

meet. Especially if students are not accustomed to talking about their writing or offering peer comments about others' work, the initial attempts at doing so may be awkward. It is important to establish clear guidelines and a safe environment if peer evaluation is to be effective.

Writing for Revision

Greenwald (1997) advocated the use of a highlighter for highlighting errors in students' papers. She maintained that the technique called students' attention to the weak areas and encouraged discussion helpful for revision.

McGovern and Hogshead (1990) designed a writing activity known as the *telescoping paper* that incorporated writing for revision into a manageable series of stages toward the perfection of a seven to ten page research paper. Students first explored a topic, producing annotations on two or three articles. The instructor provided extensive written feedback for this assignment; however, the grading weight of the assignment was very low. Students used the feedback when gathering more sources and expanding the annotations into a seven to ten page literature review. Again, instructor comments were extensive, while the weight of the assignment was low. Sections of the paper were then assigned for expansion, yet the students had to keep their papers within the original length requirements. The addition of the new sections forced the students to tighten their writing, and grading weight increased, so that students were rewarded for successful revisions. The new sections required less critical scrutiny than the earlier versions of the paper, making

the grading load manageable. Students found the editing for the telescoping length to be challenging, but they felt that the writing of the papers was more manageable than developing multiple topics into separate papers.

Formal Assessment of Writing

Grading Student Writing

Research has shown that the scoring of essays and papers is usually unreliable; scores not only vary across different graders, they vary with the individual grader at different times. Good grading practices increase the reliability of assessment judgments. Written work should be judged on its content, organization, and style. Instructors may wish to evaluate the work in each of these areas and assign a mark on the basis of some combination of these factors.

Comments should be written judiciously and legibly. Use the margins, the back, or attach a note. Try to say enough so that the student has a reasonably good chance of doing better next time; however, strive for a few analytical comments on the good and bad aspects of the work rather than a detailed critique--writing too many comments tends to overwhelm students.

Distributing a model response with the corrected essays can alleviate some of the burden of writing comments on exams. Students tend to learn a little more when they compare their answers with the model, and they develop a clearer picture of why they received the grade they did. Consider asking students producing high quality work for permission to use their response as the model. The work, sans student's name, may be used in future classes.

Suggestions for Increasing the Reliability of Ranking Methods of Evaluation

1. Read a few papers before you actually start grading in order to get an idea of the range of quality.
2. Some instructors select "range finder" papers—middle range A, B, C and D papers to which they refer for comparison.
3. Stop grading when you get too tired or bored. When you start again, read over the last couple of papers you graded to make sure you were fair.
4. Conceal the student's name while you grade the response. If you know the identity of the student, your overall impressions of that student's work will inevitably influence the scoring of the test.
5. If there is more than one essay question on the test, grade each essay separately rather than grading a student's entire test at once. Otherwise, a brilliant performance on the first question may overshadow weaker answers in other questions (or vice-versa).
6. Remain open to legitimate interpretations of the questions different from your own. If students misinterpret the intent of your question, or if your standards are unrealistically high or low, you should alter your model response in light of this information.

Two Methods for Evaluating Essay

Responses

Analytic Method

In this method the ideal or model answer is broken down into several specific points regarding content. A specific subtotal point value is assigned to each. When reading the exam, you need to decide how much of each maximum subtotal you judge the student's answer to have earned. When using this method, be sure to outline the model (ideal or

acceptable) answer BEFORE you begin to read the essays.

Holistic Method

In this method the rater reads the entire essay and makes an overall judgment about how successfully the student has covered everything that was expected in the answer and assigns the paper to a category (grade). Generally, five to nine categories are sufficient. Ideally, all of the essays should be read quickly and sorted into five to nine piles, then each pile reread to check that every essay has been accurately (fairly) assigned to that pile which will be given a specific score or letter grade.

EXAMPLE OF A DIAGNOSTIC SCALE FOR GRADING A PARAGRAPH
(Central Florida Community College)

MAIN IDEA/TOPIC SENTENCE

- 6—Presents or implies a main idea with noticeable coherence.
- 5—Presents or implies a main idea with convincing, specific detail.
- 4—Presents or implies a main idea and suggests a plan of development, which is usually carried out.
- 3—Presents or implies a main idea and suggests a plan of development, which is partially or weakly carried out.
- 2—Presents an unfocused or generalized main idea.
- 1—Presents little or no main idea, vaguely worded.

DETAILS

- 6—Are substantive, sophisticated, and elaborated.
- 5—Are fresh, mature, and extensively developed.
- 4—Are specific enough to contribute to the main idea.
- 3—Are generalized or a listing, poor support of main idea.
- 2—Are generalized, sketchy, and/or illogical.
- 1—Are very generalized, superficial, and/or rambling.

SENTENCES AND DICTION

- 6—Are varied, precise and purposeful, often polished.
- 5—Are varied and show an excellent command of language.
- 4—Are sometimes varied and show a good command of language.
- 3—Are not varied, pedestrian, and somewhat repetitious.
- 2—Are simplistic, repetitious, and sometimes disjointed.
- 1—Are tangled, incoherent, and confusing.

USAGE AND MECHANICS

- 6—Control of sentence structure, usage, and mechanics, despite an occasional flaw, visibly contributes to the writer's ability to communicate the main idea.
- 5—Control of sentence structure, usage, and mechanics, despite an occasional flaw, contributes to the writer's ability to communicate the main idea.
- 4—Occasional errors in sentence structure, usage, and mechanics do not interfere with the writer's ability to communicate the main idea.
- 3—Errors in sentence structure, usage, and mechanics sometimes interfere with the writer's ability to communicate the main idea.
- 2—Errors in sentence structure, usage, and mechanics frequently interfere with the writer's ability to communicate the main idea.
- 1—Numerous errors in sentence structure, usage, and mechanics substantially interfere with the writer's ability to communicate the main idea.

Conversion Grading Scale/Added Total Points

24 = 100	20 = 90	16 = 80	12 = 70	8 = 60
23 = 98	19 = 88	15 = 78	11 = 68	7 = 55
22 = 95	18 = 85	14 = 75	10 = 65	6 = 53
21 = 93	17 = 83	13 = 73	9 = 63	5 = 50

An Analytic Scale for Grading Content Writing

This sample scale attributes 70% of the grade to the successful explication of three content objectives, one weighted 30%, two others valued at 20%. An additional 30% of the grade is attributable to writing quality, divided equally among organization, clarity, and correctness. Space is left after each category for instructor comments.

Content objective A (30%) 2 4 6 8 10

x 3 = _____

Content objective B (20%) 2 4 6 8 10

x 2 = _____

Content objective C (20%) 2 4 6 8 10

x 2 = _____

Comments

Writing (30%)

_____ Organization (10%) 2 4 6 8 10

_____ Clarity (10%) 2 4 6 8 10

_____ Correctness (10%) 2 4 6 8 10

Comments

TOTAL _____ of 100

Overall reaction and suggestions:

from Tchudi, Stephen N. (1986). Teaching Writing in the Content Areas: College Level. NEA, p. 57.

RUBRIC FOR HOLISTIC SCORING OF ANALYSIS OF AN ARGUMENT

6 Outstanding

A 6 paper presents a cogent, well-articulated critique of the argument and demonstrates mastery of the elements of effective writing. A typical paper in this category

- clearly identifies and insightfully analyzes important features of the argument
- develops ideas cogently, organizes them logically, and connects them with clear transitions
- effectively supports the main points of the critique
- demonstrates control of language, including diction and syntactic variety
- demonstrates facility with the conventions of standard written English but may have minor flaws

5 Strong

A 5 paper presents a well-developed critique of the argument and demonstrates good control of the elements of effective writing. A typical paper in this category

- clearly identifies important features of the argument and analyzes them in a generally thoughtful way
- develops ideas clearly, organizes them logically, and connects them with appropriate transitions
- sensibly supports the main points of the critique
- demonstrates control of the language, including diction and syntactic variety
- demonstrates facility with the conventions of standard written English but may have occasional flaws

4 Adequate

A 4 paper presents a competent critique of the argument and demonstrates adequate control of the elements of writing. A typical paper in this category

- identifies and analyzes important features of the argument
- develops and organizes ideas satisfactorily but may not connect them with transitions
- supports the main points of the critique
- demonstrates sufficient control of language to convey ideas with reasonable clarity
- generally follows the conventions of standard written English but may have some flaws

3 Limited

A 3 paper demonstrates some competence in analytical writing skills and in its control of the elements of writing but is plainly flawed. A typical paper in this category exhibits one or more of the following characteristics:

- does not identify or analyze most of the important features of the argument, although some analysis of the argument is present
- devotes most of its time to analyzing tangential or irrelevant issues
- is limited in the logical development and organization of ideas
- offers support of little relevance and value for points of the critique
- does not convey meaning clearly
- contains occasional major errors or frequent minor errors in grammar, usage, and mechanics

2 Seriously Flawed

A 2 paper demonstrates serious weaknesses in analytical writing skills. A typical paper in this category exhibits one or more of the following characteristics:

- does not present a critique based on logical analysis, but may instead present the writer's own views on the subject
- does not develop ideas or is disorganized
- provides little, if any, relevant or reasonable support
- has serious and frequent problems in the use of language and in sentence structure
- contains numerous errors in grammar, usage, and mechanics that interfere with meaning

1 Fundamentally Deficient

A 1 paper demonstrates fundamental deficiencies in analytical writing skills. A typical paper in this category exhibits more than one of the following characteristics:

- provides little evidence of the ability to understand and analyze the argument
- provides little evidence of the ability to develop an organized response
- has severe and persistent errors in language and sentence structure
- contains a pervasive pattern or errors in grammar, usage, and mechanics that results in incoherence

0—illegible, off-topic, in a foreign language, or merely copies the topic
NR—blank or nonverbal

(Adapted from a holistic scoring rubric used for the GMAT exam.)

	Content	Organization	Diction	Sentence Structure	Grammar and Mechanics
20	Writer consciously shapes the introduction to establish a distinct relationship with the reader; convincingly gains reader acceptance of argument through imaginative, logical, and precise development of thesis; articulately develops and details; deliberately shapes the conclusion for convincing and persuasive appeal.	Writer designs the progression of ideas with thoughtful precision and imagination; demonstrates impressive facility in sustaining focus and establishing provocative connections for the reader to consider.	Writer chooses words with commanding sense of purpose, resulting in articulate, mature, and often compelling prose; insightful use of language; efficacious use of voice, appropriate to purpose and audience.	Writer demonstrates an impressive understanding of emphasis, rhythm, and pacing in forming and positioning sentences; often goes beyond conventions deliberately to create an effect.	Writer demonstrates a command of grammar and mechanics to create involving, often stylistic prose; control and purpose consistently evident.
18	Writer creates reader interest by introducing central idea clearly and effectively; achieves reader acceptance of argument through logical and precise development of thesis; develops and details; concludes with distinct persuasive appeal.	Writer controls development of essay by shaping a distinct beginning, middle, and end; controls thesis and progression of ideas by sustaining clear focus and consistent line of argument; organizes specific details in logical sequence; uses effective transitions to maintain coherence and provide necessary links between and within paragraphs.	Writer chooses concrete, specific words and uses them correctly; uses diction that is distinctive and mature, with effective metaphors and analogies for clarity or emphasis; avoids colloquialisms, clichés, and trite expressions; develops economical and natural style, neither wordy nor contrived or inflated; selects strong verbs with active voice predominant.	Writer understands correct use of coordination, subordination, and sentence types (simple, compound, complex, compound-complex); seeks variety in both sentence length and structural patterns.	Writer demonstrates command of mechanics; subjects and verbs agree and tenses are consistent; sentences are complete; pronouns in correct cases agree with and refer clearly to their antecedents; modifiers are properly placed; spelling and punctuation are correct. Format is correct.
16	Writer creates some reader interest in argument by presenting clear thesis statement and supporting it with good examples and reasoning. Writer's presentation of topic is not unique, yet the presentation is smooth and effective; conclusion is not strongly persuasive.	Writer controls development of essay by arranging examples supporting the thesis in an orderly and logical fashion; connects examples and reasons with adequate transitions.	Writer uses clear, concise words and relatively mature, natural style; minimizes wordiness, passive voice, and inaccurate words.	Writer demonstrates better than average variety in sentence structure; makes relatively few errors in use of coordination, subordination and sentence types (simple, compound, complex, compound-complex).	Writer exhibits occasional but limited errors in syntax, agreement, pronoun reference, spelling, or punctuation.
14	Writer frames topic in conventional and predictable manner, stating the obvious, developing only surface meaning; development may lack clarity; concludes vaguely or abruptly.	Writer employs a loose and sometimes unclear logic or pattern; needs better transitions between ideas; or adopts a mechanical development and seems to just follow a formula.	Writer uses overly general, vague, or pedestrian words; depends on clichés and jargon; overly wordy; overuses passive voice.	Writer seldom varies from simple subject-verb-complement structure; occasionally creates awkward sentences, fragments, or run-on sentences.	Writer makes frequent but manageable errors in syntax, agreement, pronoun case and reference, spelling, or punctuation. Shows problems with format.
10	Writer fails to provide a relevant discussion of the topic; does not provide evidence to substantiate an argument; does not follow a central line of discussion; commits many logical fallacies; strays from the point.	Writer demonstrates little control of topic; insufficient evidence or examples to organize	Writer chooses words almost carelessly; may be too familiar with the reader; uses colloquialisms, clichés, and jargon.	Writer structures sentences that are awkward and lack clarity; demonstrates little sentence variety; frequently creates fragments and run-ons.	Writer consistently makes basic errors in syntax, agreement, reference, spelling, or punctuation. Format is wrong or incomplete.
0	Writer fails to complete assignment.				

Grading Rubric From the University of Central Florida

A Note about Essay Exams

Many instructors consider essay questions to be the ideal form of testing, since essays seem to require more effort from the student than other types of questions. Essay questions can test complex thought processes, critical thinking, and problem-solving skills, and essays require students to use the English language to communicate in sentences and paragraphs—a skill that undergraduates need to exercise more frequently. Essay responses allow us to see our students' thought processes that lead to the answers.

While essay questions are relatively efficient to compose, the evaluation of the responses can be very time-consuming. As with essay prompts and other writing assignments, the instructor should form a model response ahead of time and clearly communicate the performance expectations. Allowing students to select which essay questions to answer (e.g. "choose two out of five") is not a good practice, as it is virtually impossible to compose five equivalent essay questions, and students will usually choose weaker questions and thereby reduce the validity of the exam.

A Checklist for Creating Essay Exam Items

- The essay item tests a higher-level learning outcome or complex content not readily measured by objective-type items.
- The item requires students to apply knowledge, integrate their learning, be creative, and demonstrate other similar skills.
- The item samples important content learned in the course.
- The item adequately evaluates the content area and level of learning intended.
- The item is written clearly.
- The item gives direction about how to respond to avoid writing all that is known about the topic.
- If extended response, the item is not too broad.
- If restricted response, the content could not be assessed more easily with an objective item.

Affective Assessment of Student Writers

The Affective Domain

Receiving = Open and attentive to new ideas
Responding = React to new information
Valuing = Apply criteria to new information
Organizing = Create schema for using information
Characterizing = Apply a belief system to new ideas

While instructors may not wish to assign a grade for affective elements of students' learning experiences, it is worthwhile to note that the informal assessment methods as prescribed in this paper should engage students in an active community of learners as they work together to learn course content and to communicate effectively in writing.

Measures of this domain could include the following:

- (1) Attitude toward the writing process as revealed in self-assessments and revision efforts,
- (2) Evidence of communication with peers in discussing writing and in providing effective peer assessments, and
- (3) Evidence and degree of reflection upon one's own writing.

REFERENCES

- Cleland, Marilyn M. (1995, December 1). Oral grammar and written grammar: The place of grammar in the developmental writing class. Innovation Abstracts, 17. The National Institute for Staff and Organizational Development.
- Dunn, Dana S. (1994, December). Lessons learned from an interdisciplinary writing course: Implications for student writing in psychology. Teaching of Psychology, 21, 223-227.
- Elbow, Peter. (1993). Ranking, evaluating, and liking: Sorting out three forms of judgment. College English, 55, 187-206.
- Greenwald, Sara. (1997, March 21). A simple way to encourage proofreading. Innovation Abstracts, 19. The National Institute for Staff and Organizational Development.
- Gruber, Sibylle; Larson, Debra; Scott, David; Neville, Melvin. (1999, Fall). Writing4Practice in engineering courses: Implementation and assessment approaches. Technical Communication Quarterly, 8, 419-440.
- IRA & NCTE Joint Task Force. (1994). Standards for the assessment of reading and writing. Urban, IL: National Council of Teachers of English.
- McGovern, Thomas V. & Hogshead, Deborah L. (1990, February). Managing writing: Learning about writing, thinking about teaching. Teaching of Psychology, 17, 5-10.
- Tchudi, Stephen N. (1986). Teaching Writing in the Content Areas: College Level. NEA.
- White, Edward M. (1994). Assigning, responding, evaluating. New York: St. Martin's Press.
- Zinn, Ava. (1998, Winter). Ideas in practice: Assessing writing in the developmental classroom. Journal of Developmental Education, 22, 28-33.

Legal Crib Sheets: Promoting Deep Levels of Processing and Learning

MaryJo O. Smith, Ph.D., Embry-Riddle Aeronautical University

S. Jay Samuels, Ph.D., University of Minnesota

Abstract

The article demonstrates that legal crib sheets encourage deep levels of information processing, an activity that should enhance students' learning and long-term memory. Levels of processing theory states that memory processes exist on a depth continuum; comprehension and synthesis are examples of deep cognitive processes that enhance memory, whereas simple repetition or examining surface characteristics of words are examples of shallow cognitive processes (Craik & Lockhart, 1972). The use of legal crib sheets allowed students to attain significantly higher overall mean test scores while not affecting their long-term retention of the material. Deeper levels of processing occurred because the students manipulated course information in preparing legal crib sheets.

Introduction

The purpose of this article is to demonstrate that students' use of legal crib sheets encourages deep levels of processing, an activity that should enhance learning and long-term memory. The levels of processing theory is a perspective that states that memory processes exist on a depth continuum. For the last quarter of a century, the Craik and Lockhart (1972) levels of processing hypothesis has had a major impact on memory theory. One of the principles underlying the theory is that the strength and durability of the memory trace can be explained as a by-product of the type of cognitive processing used to input information into long-term memory. Comprehension, categorization, and synthesis were thought to be deep cognitive processes that enhance memory. Simple repetition or examining a surface characteristic of a word, such as, does "snails" have one or two syllables, were thought to be shallow cognitive processes that did not enhance memory. From their research, Craik and Lockhart concluded that memory performance is strongly linked to the type of processing used to store information.

The 1969 Hyde and Jenkins memory study was the precursor to Craik and Lockhart's (1972) groundbreaking work on memory theory and memory enhancement. In the Hyde and Jenkins

study, four groups of students were given the task of remembering 12 word-pairs (24 words) that were presented in random order. Group 1 was given an intentional learning task; they were told to memorize the word-pairs because they would have to recall them at a later point in the procedure. Groups 2, 3, and 4 were incidental learning groups. These groups were not told beforehand that they would be required to recall the word pairs; each group was asked to make a different kind of judgmental decision about the words. The different kinds of judgmental decisions were designed to produce different levels of processing. In the experiment, Group 2 was asked to look at each word and decide if the word had letter "e" in it. Group 3 was asked to count how many letters were in each word as it was presented. Group 4 was asked to look at each word and decide if it was pleasant or unpleasant. The processing levels for the incidental learning groups ranged from shallow (Groups 2 and 3) to deep processing (Group 4). Shallow processing directed attention away from the meaning of the word to a surface characteristic such as its spelling or length. Deep processing, on the other hand, directed attention to a semantic characteristic of the word such as its affective dimension, i.e., pleasant-unpleasant. The most striking finding in this study was that Group 4, which used deep

processing to evaluate words on their pleasant-unpleasant dimension, recalled as many words (16.1) as the students in the intentional-learning group (16.3) who were told to memorize the words. Those students who were in the other two incidental conditions (Groups 2 and 3), where processing was based on surface characteristics of the words such as the presence of letter "e" or word length, recalled only 9.4 and 9.9 words respectively.

Craik and Lockhart (1972), working from Hyde and Jenkins' (1969) findings, developed the levels of processing theory as an alternative to a simple information processing model of memory that contained components such as sensory information store, short-term memory, and long-term memory. Each memory component has different characteristics for the variables: memory capacity, rates of decay, speed of input, and speed of output. Craik and Lockhart suggested that the determinant of how much information was to be stored and how long it was to be remembered was not where the information was stored in memory, e.g., short-term or long-term. Instead, they argued that memory storage was determined by the type of encoding process used to input the information. If shallow information procedures were used, the ability to remember the information was not as good as the recall produced by deep processing. Lockhart and Craik's more recent work in 1979 and 1990 has sought a

rapprochement between the levels of processing and information processing models of memory and has attempted to operationalize the encoding strategies that facilitate deep levels of processing.

Memory trace is now generally accepted to be a by-product of cognitive processes such as comprehending, categorizing, conceptualizing, synthesizing, and elaborating (Craik & Lockhart, 1972). Kiewra (1983) reviewed the research on note-taking. He concluded that the act of note-taking by itself is beneficial, independent of the reviewing process that could be done on the notes. What made the process of note-taking beneficial was the extent to which the student was able to make the new information meaningful. In summary, the levels of processing theory stated that more learning and retention will occur when: 1) people work harder at encoding, 2) the information that is to be learned is related to and compared with information that is already in memory, and 3) the approach to learning new information emphasizes constructing meaning (Craik & Lockhart).

Hypothesis

In this experiment, we were testing the hypothesis that when students compress and synthesize a large body of information from their text and lectures in order to construct a legal crib sheet, that deep processing of information is unavoidable. Therefore, when students construct legal

crib sheets, the process should enhance their learning and retention of the course material. Before each exam, the students were given an oral review of the lectures and a written list of the important concepts from the book. The students did not know exactly which concepts would be tested; therefore, what was written on their crib sheets resulted from their thinking about what information was important and likely to be tested. These activities can foster deep levels of processing.

Method

Subjects

The subjects were students in the Learning, Cognition, and Assessment in Schools course, an introductory Educational Psychology course in the Master's of Education program. The course was required for their teacher certification. For these students, the ratio of females to males was 2 to 1; their ethnicity was primarily white; the average age was 31; and on average, they maintained a 3.13 GPA.

The Educational Psychology course in which the research was done was taught during each of the fall and winter quarters of the 1997-98 academic year at the University of Minnesota. The fall enrollment was 108 students and the winter enrollment was 76 students; a few students took the exams in accordance with the University disability policy and their scores could not be used in this analysis. Before this study, statistical comparisons of test grades for the fall and

winter classes indicated that there were no differences between previous classes. The students attended class for four hours per week: three hours of lecture with the professor and one hour of lab in classes of approximately 30 with a graduate student lab instructor. The students in the fall 1997 class served as the experimental group; the students in the winter 1998 class served as the control group. Because random assignment of students to the treatment groups was not possible, this convenience sample was used.

Two exams were given in the course. Students in the fall course were permitted to bring one piece of 8½" x 11" paper to use as a crib sheet during the exams; students in the winter course were not permitted to use a legal crib sheet.

Procedures

Identical instruction, including lectures, reading assignments, and small group activities, were given to both classes. The students were given two multiple-choice exams, each containing 45-questions; identical exams were used during the fall and winter quarter. Exam 1 was given after the sixth class in the quarter; Exam 2 was given after the thirteenth class in the quarter. Each exam covered the content area from the lectures, lab, and book; neither exam was cumulative. The students were given 50 minutes to complete the exams.

At the conclusion of Exam 1 for the fall 1997 quarter, the students were asked to

reflect on their use of their crib sheet.

Eighty-seven students responded to these four open-ended questions:

1. What strategies did you use when formulating your crib sheet?
2. What portion of your crib sheet was from the lectures? From the text?
3. How much did you use the crib sheet during the exam?
4. Did you find the crib sheet to be helpful? How? Why?

Results and Discussion

The descriptive statistics, Table 1, describe the scores on Exam 1 and Exam 2 for both classes. The students taking the class in fall 1997, who had made legal crib sheets, had a higher mean for Exam 1 than the students taking the class during winter 1998, who did not have crib sheets. The descriptive statistics for Exam 2 again showed that the students taking the class in fall 1997, who had made legal crib sheets, had a higher mean than the students taking the class in winter 1998.

Table 1

Descriptive Statistics

	Used Legal Crib Sheets		Did Not Use Legal Crib Sheet	
	Fall Exam 1	Fall Exam 2	Winter Exam 1	Winter Exam 2
N	104	102	74	76
Mean	37.61	37.23	36.39	34.92
Std. Deviation	3.42	4.08	4.21	3.81
Range	24.00	30.00	23.00	20.00
Minimum	19.00	13.00	19.00	22.00
Maximum	43.00	43.00	42.00	42.00

Figure 1. Summary of the Comparisons Tested.

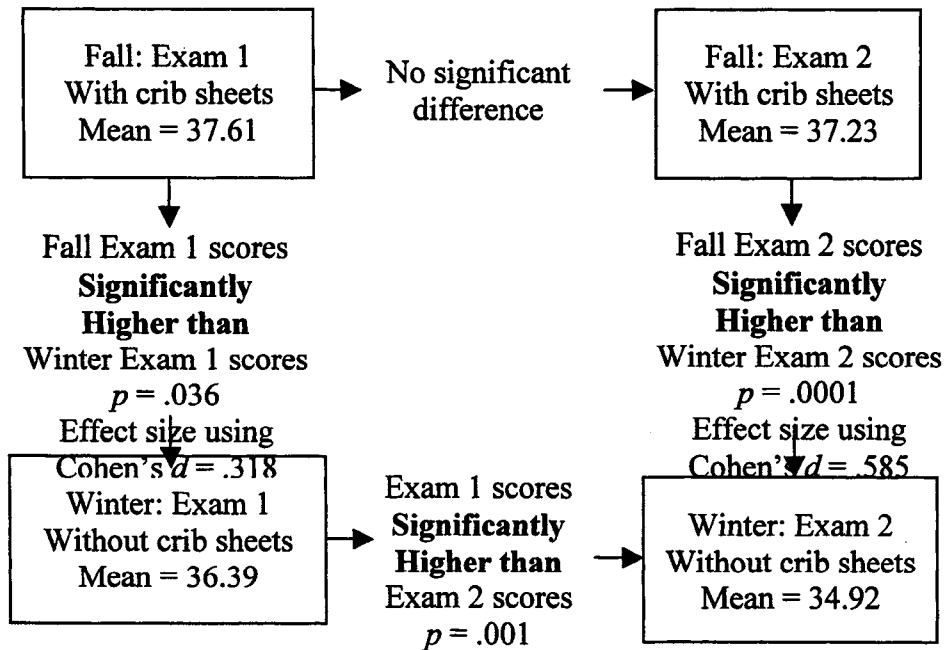


Figure 1 summarizes the experimental combinations that were used in the study. Paired-sampled t-tests were run for each individual class: fall 1997 and winter 1998. There was no statistical difference ($t=.809$, $df=101$, $p=.42$) between the students' overall paired Exam 1 and Exam 2 scores taken during the fall 1997. Therefore, the students' scores were similar for both exams and the students were able to maintain a high score for both exams because crib sheets were allowed on both exams. For the winter 1998 students, who did not use crib sheets, their overall Exam 1 scores were significantly higher ($t=3.335$, $df=71$, $p=.001$) than their Exam 2 scores. Therefore, it appears that a crib sheet might have been helpful in allowing these students to maintain their same grade level.

The results of the independent samples t-test for the Exam 1 showed that the fall 1997 students, who were permitted to use crib sheets, scored significantly higher than the winter 1998 students who did not use crib sheets ($t=2.118$, $df=176$,

$p=.036$), as shown in Figure 1. The treatment effect, as measured by the Cohen's d , was .318, which indicated a small to medium effect size. An effect size of .3 indicates that the mean score of the students using legal crib sheets is at the 62nd percentile of the students who were not permitted to use crib sheets (Becker, 1998). Therefore, this analysis supports the hypothesis that crib sheets promote deep levels of processing resulting in higher levels of learning as expressed in test scores.

The results of the independent samples t-test for Exam 2 showed that there was a significant difference between the scores for students who used legal crib sheets and those who did not ($t=3.666$, $df=172$, $p=.0001$). The treatment effect, as measured by the Cohen's d , was .585, which indicated a medium effect size. An effect size of .6 indicates that the mean score of the students using legal crib sheets is at the 73rd percentile of the students who were not permitted to use crib sheets (Becker, 1998). The students who were

permitted to use the crib sheets scored significantly higher than the students who were not permitted to use a crib sheet.

Qualitative Results

Eighty-seven students responded to the qualitative questions asked after Exam 1 in the fall of 1997. The students' listed the following strategies for how they formulated their crib sheets: using the information from the review session, defining key words and concepts, writing dates for events mentioned in class and in the text, writing notes on important people or articles specifically mentioned during lectures, and including information that they were still learning and had yet to master.

Thirty-two out of 87 students (37%) responded that half of their crib sheet was from lecture notes and half was from the text. Twenty-one students (24%) responded that approximately $\frac{1}{2}$ of their crib sheet was from the lecture and that $\frac{1}{2}$ was from the text. Seventeen students (20%) responded that approximately $\frac{1}{4}$ or less of their crib sheet was from the lecture and $\frac{3}{4}$ was from the text. Eleven students (13%) responded that $\frac{1}{2}$ or more of their crib sheet had more

information from the lecture and $\frac{1}{2}$ or less contained information from the text. Six students (6%) did not respond to the question.

Twenty-nine out of 87 students (33%) said that they used the crib sheet very little. Many students were surprised at how little they used the crib sheet. One wrote, "Less than what I thought I would have to – I guess I learned a lot just doing the crib sheet!" Another wrote, "Not as much as I thought I would. Writing the crib sheet was a good way to study." Finally, "I hardly used it at all during the exam. When I did, it was just to check an answer I had already given." Thirty-four of the 87 (39%) students used the crib sheet for five questions or less. Twenty-three (26%) students stated that the crib sheet was truly helpful when they took the exam; they used their crib sheets for 25% - 50% of the questions. One student did not respond.

Sixty-eight out of 87 students (78%) responded that they felt that the crib sheet was helpful. Only six students (7%) responded that they felt the crib sheet was

not helpful. Thirteen students (15%) did not respond to the questions.

Fifty-one of 87 students (59%) explained how or why the crib sheet was helpful. Forty-one students responded that the process of creating the crib sheet was most helpful. "While it was helpful to have the crib sheet on hand (for the 20% of the questions I used it for), I really learned the most simply by making this crib sheet. It's a great idea." Another student wrote, "I found the crib sheet helpful because it provided me with a great study opportunity. I feel that is why I didn't need to use it very much because I really got to learn the material." Lastly, a student wrote, "I thought this was a good learning tool because I learned a lot while I wrote out my crib sheet. I probably learned more making my crib sheet than I would have learned just studying." Five students reported that the crib sheet helped minimize or eliminate their test anxiety. Five students also reported that the crib sheet was a confidence booster or security blanket.

Conclusion

Legal crib sheets allowed the students to obtain significantly higher overall mean test scores while not affecting their long-term retention of the material. Therefore, deeper levels of processing occurred because the students were manipulating the information as they decided how to represent the course information on their crib sheets. The majority of the students did not extensively use the crib sheet during the exam because they had cognitively learned the material. Depth of processing suggests that the students were able to synthesize the text and lecture materials more deeply while creating the crib sheet. Additionally, they were engaging in good study habits. The qualitative information provided by the students indicated that they appreciated having the crib sheets.

A construct that was not asked qualitatively, but that some student addressed in their qualitative responses was test anxiety. For those five students who

Legal Crib Sheets

normally have test anxiety, the ability to use a legal crib sheet greatly reduced their anxiety level and allowed them to represent their knowledge more accurately.

Our Educational Psychology course will continue to use legal crib sheets for several reasons: the higher test scores, the reduction in their test anxiety, and the help in organizing their learning. The professors appreciated the deeper processing of the course materials that the legalized crib sheet provided. The creation of the legalized crib sheet is another tool to help our students learn the course material.

References

- Becker, L. A. (1998). Effect size (ES). Available: <http://www.uccs.edu/~lbecker/psy590/es.htm> [2001, November 23].
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11(6), 671-684.
- Hyde, T. S. & Jenkins, J. J. (1969). The differential effects of incidental tasks on the organization of recall of a list of highly associated words. Journal of Experimental Psychology, 82, 472-481.
- Kiewra, K. A. (1983). The process of review: A levels of processing approach. Contemporary Educational Psychology, 8(4), 366-374.
- Lockhart, R. S. & Craik, F. I. M. (1979). Levels of processing: A reply to Eysenck. British Journal of Psychology, 69(2), 171-175.
- Lockhart, R. S. & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary in a framework for memory research. Canadian Journal of Psychology, 44(1), 87-112.

**COGNITIVE AND AFFECTIVE DOMAIN
LEARNING ASSESSMENT CHOICES**

by

Ronald Clark
Joseph Price

ACKNOWLEDGEMENTS

The researchers wish to acknowledge the contribution of their colleagues and students who volunteered to be subjects in this research study. They also wish to acknowledge the able assistance of Kathleen Citro of the Hunt Memorial Library staff of ERAU for her timely and comprehensive reference material search support. They acknowledge the research leadership role played by the ERAU Faculty Development Committee and its chair, Dr. Earl Wheeler. Last, they acknowledge the contribution of their families for their support of time and encouragement.

Through the research and development of this paper, the researchers have gained valuable insights to augment and modify their personal student learning assessment schemas, and to provide leadership and mentorship for the faculty whom they supervise and work with.

ABSTRACT

Assessing university student learning is both an academic skill and an art form, with Bloom's Taxonomy of the cognitive domain perhaps the preeminent schema in use today. This research study sought to find out if Embry-Riddle faculty and students were aware of Bloom's affective domain, and to assess the degree of satisfaction with current student learning assessment. Using a descriptive research model, 61 faculty and students were surveyed and three classes were provided with an open model of assessment. The results indicated that both faculty and students were satisfied with ERAU student learning assessment, learned more productively with student-decided assessments, and knew far less about the affective domain. It was concluded that the research should be expanded, the survey instrument should be reworked, and faculty should receive learning assessment training.

INTRODUCTION

Background of the Problem

The assessment of university student learning outcome achievement is both an art and a science, with judicious applications of both necessary to achieve a "true" evaluation. At the end of the assessment process, both students and faculty should feel that the process has been fair and accurate. After a time, faculty develop their own schema of student learning, generally with little or no student input into their own grading formula. Both scientific and artful help exists in the education assessment literature and in the halls of academe, where faculty orientation manuals and such exist, especially to assist the newer faculty member.

The "taxonomy of educational objectives of Benjamin Bloom" is widely thought to consist of only the "cognitive" categories of knowledge, comprehension, application, analysis, synthesis, and evaluation. Many references allude to "Bloom's Taxonomy" as a cognitive taxonomy, when, in fact, an affective domain exists as well (major categories, 2002). Could the apparent lack of information and understanding regarding the affective domain of "Bloom's Taxonomy" result in a lack of Embry-Riddle Aeronautical University student grade assessment along affective domain lines? Would both faculty and

students feel that student grade assessment is only practically done along more well-known cognitive lines? Is student assessment along affective domain categories practical and valid? Curious to obtain the answers to these and other similar student grade assessment questions, the researchers chose to conduct an original research study lasting six months, from January to June of 2002. They concentrated on a faculty and student sample from the Southwest Region of Embry-Riddle's Extended Campus.

Researchers' Work Settings and Roles

Doctor Ronald Clark is an Associate Professor of Aeronautical Science and a Regional Faculty Advisor (RFA) for the Southwest Region of Embry-Riddle's Extended Campus. He holds degrees in psychology, counseling and human development. He has been a college teacher since 1977, and has taught at community colleges, universities, and internationally. Since 1987, he has been a college professor for Embry-Riddle, teaching primarily at the graduate level. Since 1990, he has authored original research studies in adult learning theory, educational technology use in the classroom, and teaching basic life skills such as critical thinking, computing, speaking and writing.

Joseph (Jay) Price is the Center Academic Advisor (CAA) and Center Full

Time Faculty (CFTF) member at the Southwest Region's Las Vegas Center. He has degrees in Psychology and Guidance and Counseling. Since 1994, he has been teaching college courses for Embry-Riddle and has served as a Center Academic Advisor since 1996. Jay teaches human factors and Crew Resource Management (CRM) training for airline and armed forces flight crew across America.

Statement of the Problem

The assessment of Embry-Riddle university student achievement of learning outcomes, course by course, is most probably being accomplished along the lines of only the cognitive domain of "Bloom's taxonomy". Affective domain assessment may be indicated by faculty and students. There may be student dissatisfaction with faculty-decided (no student participation) assessment components, and a "one size fits all" mentality may not be as effective and fair as more individualized assessment.

Limitations and Assumptions

Because of a lack of funding support for this research study, the sample size for both faculty and student samples was limited to n=20 and n=41, respectively. Additionally, the timeframe for data collection was limited to two consecutive ERAU Extended Campus terms of nine weeks each, or an overall total of five months. The geographical dispersion of the

researchers between Las Vegas and Phoenix was somewhat helped by both researchers teaching in Las Vegas during the Spring II term, from March through May of 2002.

REVIEW OF RELEVANT LITERATURE AND RESEARCH

Assessment of Student Learning

According to Maki (2002), higher education institutions all too often view the assessment of student learning as a periodic activity, or compliance approach, driven by perhaps an impending accreditation visit. She contrasts this motive with that of institutional curiosity, which seeks to know which, how, what, when, students learn, and through which pedagogy and andragogy schemas. To assist institutions of higher learning in their student learning assessment planning, she developed an assessment guide that helps integrate assessment into institutional culture. Over time, the assessment of student learning is seen as becoming systematic and a part of organizational practice.

The American Association of Higher Education (AAHE) (2002) has formulated what they call nine principles of good practice for assessing student learning:

1. The assessment of student learning begins with educational values.
2. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.
3. Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.
4. Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes.
5. Assessment works best when it is ongoing not episodic.
6. Assessment fosters wider improvement when representatives from across the educational community are involved.
7. Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.
8. Assessment is most likely to lead to improvement when it is a part of a larger set of conditions that promote change.
9. Through assessment, educators meet responsibilities to students and to the public (AAHE, 2002, pp. 1-2)

The National Center for Fair & Open Testing (NCFOT) (2002), through their National Forum on Assessment, has published what they call the principles and indicators for student assessment systems, a seven step guide to the assessment of student learning:

1. The primary purpose of assessment is to improve student learning.
2. Assessment for other purposes supports student learning.
3. Assessment systems are fair to all students.
4. Professional collaboration and development support assessment.
5. The broad community participates in assessment development.
6. Communication about assessment is regular and clear
7. Assessment systems are regularly reviewed and improved (NCFOT, 2002, p. 1)

Anderson (2001) believes that the assessment of student learning should be tailored to student learning styles. He characterizes learning styles as to how we prefer to learn, specifically as to:

1. The type of information we receive (sensory vs. intuitive).
2. How we perceive information (visual vs. verbal).

3. How we organize information (inductive vs. deductive).
4. How we process information (actively vs. reflectively).
5. How we understand information (sequentially vs. globally).

Anderson goes on to classify the many dimensions of learning styles as:

reflective vs. impulsive
non-affective vs. affective
elaborative vs. shallow (repetitive)
processing
scanning (visual) vs. focusing
field-independent vs. field-sensitive
analytical vs. relational
independent vs. dependent
participant vs. avoidant (Anderson, 2001, pp. 1-2)

He sees that learning styles are not bipolar clusters, but rather continuums, wherein learners are so much of this and so much of that, along individual learning style preferences. He cautions that educators should not force students to change their learning styles to adapt to assessment schemas, but, rather, that this happen the other way around.

In arguing for fair assessment practices, Suskie (2000) states that educators make their assessments and how they use the results of assessment as fair as possible for as many students as possible. Her call is for giving students equitable opportunities to demonstrate what they know. She lists what

she calls "seven steps to fair assessment" as follows:

1. Have clearly stated learning outcomes and share them with your students.
2. Match your assessment to what you teach and vice versa.
3. Use many different measures and many different kinds of measures.
4. Help students learn how to do the assessment task.
5. Engage and encourage your students.
6. Interpret assessment rules appropriately.
7. Evaluate the outcomes of your assessments (Suskie, 2000, pp. 1-2)

Mislevy, Steinberg, and Almond (2001) argue that advances in cognitive psychology and technology make it possible to improve educational assessment. They see more complex learning assessments through the use of simulation, interactivity, collaboration and constructed response techniques. In their "evidence-centered" assessment design, learning situations and students are analyzed with databasing technology, using an advanced cognitive psychology model.

Bloom's Taxonomy: Cognitive Domain

In 1948, a distinguished group of education testing psychologists, led by Benjamin Bloom, departed the American Psychological Association (APA) national convention with both a dissatisfaction with

the current state of the art of educational testing and assessment, and an excitement regarding their input to change this. Their subsequent collaboration over the next several years led to the development of what has become widely known as "Bloom's taxonomy", a comprehensive index of educational goals or outcomes (Bloom, Englehart, Furst, Hill & Krathwohl, 1956). While three domains (cognitive, affective, and psychomotor) were devised, only the first, or cognitive, domain, published in 1956, has received widespread acceptance and use.

Bloom's Taxonomy: Affective Domain

Following the popularity of the first Bloom, et al handbook in 1956, Krathwohl, Bloom and Masia (1964) published the second handbook of series: the affective domain. According to the authors, they were interested in assessing such things as student's "interests, attitudes, appreciations, values and emotional sets or biases" (p. 7). Their affective domain consists of five levels:

- 1.0: Receiving (attending)
 - 1.1: Awareness
 - 1.2: Willingness to receive
 - 1.3: Controlled or selected attention
- 2.0: Responding
 - 2.1: Acquiescence in responding
 - 2.2: Willingness to respond
 - 2.3: Satisfaction in response
- 3.0: Valuing
 - 3.1: Acceptance of a value

- 3.2: Preference for a value
- 3.3: Commitment
- 4.0: Organization
 - 4.1: Conceptualization of a value
 - 4.2: Organization of a value system
- 5.0: Characterization by a value or value complex
 - 5.1: Generalized set
 - 5.2: Characterization (pp. 176-185)

As can be seen from the above list of affective descriptors, these are not commonly used words or assessment categories of current day educational assessment. As this research study will demonstrate, both faculty and student subjects did not really understand the words of the "Bloom's Taxonomy" affective domain, much less the domain itself.

Statement of the Research Questions

Are the faculty and student learning assessment preferences in the Southwest Region of Embry-Riddle Aeronautical University's Extended Campus the same or different? Are faculty and students comfortable with current student learning assessment practices? Do faculty and students understand (and prefer) the learning assessment categories of the Bloom's taxonomy affective domain?

RESEARCH METHODOLOGY

Research Design

The researchers decided on a descriptive model for this research project. Their assessment consisted of three parts: opening three undergraduate and graduate

Cognitive and Affective Domain
Learning Assessment Choices

What follows is a breakout of "old" and "new" course grading criteria and percentages:

MAS 515: Las Vegas Center: 13 graduate students

Old grading criteria:		New grading criteria:	
Research paper:	25%	Research paper:	30%
PPT presentation	15%	PPT presentation:	30%
Case Study:	15%	Project presentation:	30%
In class work:	10%	Class participation:	10%
Class participation:	10%		
Final exam:	25%		

This class seemed to enjoy their participation in the grading category and percentage decision. They seemed to put more effort into this course. They and the researcher felt that the class learned more.

MAS 605: Las Vegas Center: 12 graduate students

Old grading criteria:		New grading criteria:	
GRP Proposal:	50%	GRP Proposal:	100%
Open book take home final exam:	30%		
PPT presentation:	10%		
Class participation:	10%		

The researcher was surprised that the class chose 100% of their grade for the GRP Proposal, and had to administer several "no grade" descriptive and inferential statistics quizzes to augment his assessment, since the GRP Proposal does not contain any statistical applications. All of the GRP Proposals were turned in on time, with, in the researcher's estimation, an overall superior product. As one of the graduate students was influential in steering the 100% choice, he became the unnamed class leader, and the class environment and attitude was altered for the good in a very positive way.

MAS 604: Tucson Center: eight graduate students

Old grading criteria:		New grading criteria:	
Take home final exam:	25%	Take home final exam:	25%
Research paper:	60%	Research paper:	40%
PPT presentation:	10%	PPT presentation:	15%
Class participation:	5%	Class participation:	15%
		Current events presentation:	5%

Following the first class, there was a noticeable student empowerment evident. Current events presentation assignments were made and carried out well. It was apparent that the empowerment of the graduate students to choose their own grading criteria had a strong positive effect on the class.

Faculty Results

The 20 faculty surveyed indicated that they evaluated student learning using the following assessment tools and grade percentages:

(read: assessment tool: #/20: mean: SD: range (R))

Individually authored research paper:	1/20	m=25.45	SD= 8.5	R=10-40
Jointly authored research paper:	3/20	m=21.67	SD=14.4	R= 5-30
Oral final exam:	2/20	m=20	SD=0	R=10-30
Case study	8/20	m=20.62	SD=10.8	R= 5-35
Take home open book final exam:	9/20	m=27.2	SD=7.12	R=20-40
In class closed book midterm exam:	7/20	m=21.1	SD=8.6	R=10-30
In Class open book midterm exam:	7/20	m=22.9	SD=5.7	R=15-30
PowerPoint presentation of paper:	10/20	m=15.1	SD=7.07	R= 5-30
Verbal presentation of paper:	7/20	m=12.9	SD=6.36	R= 5-25
In class quizzes:	9/20	m=22	SD=18.46	R= 5-60
Other assessments:	(27)	m=18.14		R= 7-40

- Article reviews/participation/homework
- Class participation (6)
- Closed book final (2)
- Current assignment
- Current topics
- Group case study
- Group oral presentation
- Hands on practice project
- Homework
- In class closed book final (4)
- Lab demos
- Multimedia (not only PPT) presentation
- Oral presentation
- Presentation of project
- Project paper
- Take home midterm
- Tech demonstration
- Verbal debate

Cognitive and Affective Domain
Learning Assessment Choices

The following Likert Scale items were answered by the faculty as indicated:

12. Under the current ERAU academic rules, I can accurately evaluate all of my ERAU students.

Seventeen of 20 faculty answered item #12, with a mean response of 2.76, a SD of 1.89 and a range of 1-7.

13. Students can evaluate each other better than faculty can.

Seventeen of 20 faculty answered item #13, with a mean response of 5.117, a SD of 1.8, and a range of 1-7.

14. My ERAU course grades have been based on my students' awareness and attention during class.

Sixteen of 20 faculty answered item #14, with a mean response of 3.06, a SD of 1.12, and a range of 1-5.

15. My ERAU course grades have been based on my students' responding to instruction in class.

Sixteen of 20 faculty answered item #15, with a mean response of 3.16, a SD of 1.18, and a range of 1-5.

16. My ERAU course grades have been based on my students' value choices during class.

Fifteen of 20 faculty answered item #16, with a mean response of 4.47, a SD of 1.85, and a range of 2-7.

17. My ERAU course grades have been based on my students' organization of a value system during the course.

Fifteen of 20 faculty answered item #17, with a mean response of 4.6, a SD of 1.88, and a range of 2-7.

18. My ERAU course grades have been based on my students' development of value complexes in class.

Fifteen of 20 faculty answered item #18, with a mean response of 4.73, a SD of 1.83, and a range of 2-7.

In response to faculty survey item # 19: The single most correct part of my average ERAU course evaluation is the evaluation of the student's: _____, the 17 faculty responses were as follows:

Ability to logically analyze problems and choose an appropriate solution method

Ability to think as a decision-maker

Comprehension of new material

Demonstrated ability to do the course work

Define, analyze, decide and present

Exams

Grasp of concepts and procedures

Knowledge of the course material

Knowledge of the learning objectives

Learning and application

Objective knowledge

Opinion of the course value and instructor's ability to get the material across understandably

Perception of the presented material and application to their day-to-day endeavors
 Show an understanding of meteorological concepts
 Synthesis and application
 Understanding of how to prepare for the FAA written exam
 Understanding of the subject matter

Faculty survey item # 20 concerned the faculty academic evaluation of students at ERAU based on which of the following concepts that faculty felt they displayed in the classroom? (circle all that apply).

- Analysis
- Application
- Characterization by a value or value complex
- Comprehension
- Evaluation
- Knowledge
- Organization
- Receiving
- Responding
- Synthesis
- Valuing

Item # 20 was answered by the faculty as follows:

Analysis:	16 yes	1 no
Application:	15 yes	2 no
Characterization by a value or value complex:	3 yes	14 no
Comprehension:	17 yes	0 no
Evaluation:	10 yes	7 no
Knowledge :	13 yes	4 no
Organization:	10 yes	7 no
Receiving:	3 yes	14 no
Responding:	10 yes	7 no
Synthesis:	10 yes	7 no
Valuing:	15 yes	2 no

21 comments can be found in Appendix C: Faculty Data.

Student Results

The 41 students surveyed indicated that they preferred to be evaluated with the following assessment tools and grade percentages:

(read: assessment tool: #/20: mean: SD: range (R))

Individually authored research paper:	38/41	m=38.02	SD= 20.45	R=10-100
Jointly authored research paper:	18/41	m=22.22	SD=12.27	R=10-40
Case study	16/41	m=19.69	SD= 9.91	R= 5-40
Oral final exam:	14/41	m=18.21	SD= 8.23	R=10-35
Take home open book final exam:	30/41	m=25	SD= 15.20	R=10-70

Cognitive and Affective Domain
Learning Assessment Choices

In class closed book midterm exam:	13/41	m=22.7	SD= 11.43	R= 5-40
In Class open book midterm exam:	21/41	m=24.29	SD=13.72	R= 5-70
PowerPoint presentation of paper:	34/41	m=19.85	SD=11.96	R= 5-60
Verbal presentation of paper:	23/41	m=17.39	SD= 1.83	R= 5-40
In class quizzes:	10/41	m=20	SD=10.8	R= 10-40
Other assessments:	(17)	m=13.44		R= 5-40

attendance (2)
class participation (5)
class participation/homework
class subject PPT briefing
closed book final exam
current events
final exam (2)
homework
participation
weekly class project
weekly current event topics

The following Likert Scale items were answered by the faculty as indicated:

10. I am academically evaluated fairly at ERAU.

All 41 students answered item #10, with a mean response of 1.9, a SD of 1.20, and a range of 1-7.

11. Students can evaluate each other better than faculty can.

All 41 students answered item #11, with a mean response of 1.95, a SD of 1.20, and a range of 1-7.

12. My ERAU course grades have been based on my awareness and attention during class.

Forty students answered item #12, with a mean response of 2.65, a SD of 1.25, and a range of 1-6.

13. My ERAU course grades have been based on my responding to instruction in class.

Forty students answered item #13, with a mean response of 2.63, a SD of 1.23, and a range of 1-6.

14. My ERAU course grades have been based on my value choices during class.

Forty students answered item #14, with a mean response of 3.41, a SD of 1.8, and a range of 1-7.

15. My ERAU course grades have been based on my organization of a value system during the course.
Forty students answered item #15, with a mean response of 3.43, a SD of 1.69, and a range of 1-7.

16. My ERAU course grades have been based on my development of value complexes in class.
Forty students answered item #16, with a mean response of 3.21, a SD of 1.48, and a range of 1-7.

In response to item # 17: The single most correct part of my average ERAU academic course evaluation is the evaluation of my : _____,

the 34 student responses were as follows:

application
attendance
communication skills (2)
development and value
GRP
knowledge (2)
meeting course objectives--learning the material
knowledge of course concepts
paper (4)
paper/briefs/test
paper with presentation (2)
participation (2)
presentation/research
research
research projects
responsiveness to the teacher's teaching methods
tests and research papers
test scores (2)
the effort I put into each class
the quality of material I present or turn in to class
work (2)
work completed
writing
writing skills

Student survey item # 18: My academic evaluation at ERAU has been based on which of the following concepts that I displayed in the classroom?
(circle all that apply)

Analysis
Application
Characterization by a value or value complex
Comprehension

Cognitive and Affective Domain
Learning Assessment Choices

Evaluation
Knowledge
Organization
Receiving
Responding
Synthesis
Valuing

Item # 18 was answered by the students as follows:

Analysis:		26 yes	15 no
Application:		28 yes	13 no
Characterization by a value or value complex:	2 yes		39 no
Comprehension:			
	30 yes		11 no
Evaluation:			
	12 yes		29 no
Knowledge :			
	31 yes		10 no
Organization:			
	14 yes		27 no
Receiving:			
	9 yes		32 no
Responding:			
	21 yes		20 no
Synthesis:			
	5 yes		36 no
Valuing:			
	8 yes		33 no

Item # 19 comments can be found in
Appendix D: Student Data.

DISCUSSION

Re-assessing Course Grading Through Student Choice

In all three graduate classes, the student choice of grading criteria had a noticeable and positive effect on the overall class environment and in the quality and timeliness of the class work produced. Class leaders emerged and assisted the class in a positive way. It was apparent that the student buy-in for their own assessment was a powerful academic tool, and one which the researcher's intend to use in the future.

Faculty Results

The majority of the faculty surveyed indicated that they had not received training regarding the assessment of student learning. It was clear that several faculty misunderstood the intended use of the word "evaluation, so a skew exists in these results. Seven faculty indicated that other schools had "better" student learning assessment techniques or practices. Most agreed that ERAU has fair student learning assessment practices.

It appears that the faculty used a wide variety of student learning assessment techniques, with a variable percentage of the students' grades spread among several assessment techniques. Take home open book final exams received the largest grading percentage at 27.2%, followed by individually authored research papers at 25.45%, in-class open book midterm exams

at 22.9%, and in-class quizzes at 22%. Not all faculty reported using all of the above listed assessment tools, so the data are skewed.

The faculty generally agreed that they can, under the current ERAU academic rules, accurately evaluate their students. They disagreed that students can evaluate themselves better than faculty can. On the affective Likert Scale items, the faculty generally agreed with their assessments based on "awareness" and "attention", but disagreed that they assess grades based upon students' "value choices", "organization of a value system", and "development of value complexes".

When asked about their academic evaluation of students based upon a mixture of Bloom's Taxonomy cognitive and affective domain key level words, they responded with 81% "yes" responses to cognitive domain key words, as compared to 40% "yes" responses to affective domain key words. While this is considered a significant difference, and a key finding of this research study, there appears to be a lack of understanding among the faculty as to affective domain level meaning.

Student Results

Only four of 41 students surveyed indicated that they had received "better" assessments of their academic learning than at ERAU. This is considered a significant research finding. It appears that the

students preferred a wide variety of student learning assessment techniques, with a variable percentage of their grades spread among several assessment techniques. Individually authored research papers received the largest grading percentage at 38%, followed by take home open book final exams at 25%, in-class open book midterm exams at 24.3%, in-class closed book midterm exams at 22.7%, jointly authored research papers at 22.2%, and verbal presentations of a paper at 20%.

The majority of the students surveyed felt that they were academically evaluated fairly at ERAU. Surprisingly, they strongly indicated that they could evaluate other students better than faculty can. The students somewhat agreed that they have been evaluated on their "awareness" and "attention", "responding to instruction", "value choices", "organization of a value system", and "development of value complexes". Their aggregate indications of affective domain evaluation, although weak at 3.065 on a Likert Scale of 7 choices, where "1" is "completely agree", are surprising, and may be due to misunderstanding, rather than positive choice.

The faculty and students differed somewhat in their choice of their core academic learning assessment "trait" upon which to be evaluated as is seen below:

Faculty	Students
Application	Responsiveness to the
Attendance	teachers's teaching methods
Communication skills (2)	Tests and research papers
Development and value	Test Scores (2)
GRP	The effort I put into each class
Knowledge (2)	The quality of material I present
Meeting course objectives—learning	or turn in to class
the material	Work (2)
Knowledge of course concepts	Work completed
Paper (4)	Writing
Paper/briefs/test	Ability to logically analyze
Paper with presentation (2)	problems and choose an
Participation (2)	appropriate solution method
Presentation/research	Ability to think as a decision
Research	maker
Research projects	Comprehension of new material
Demonstrated ability to do the coursework	Perception of the presented
Define, analyze, decide and present	material and application to their
Exams	day-to-day endeavors
Grasp of concepts and procedures	Show an understanding of
Knowledge of the course material	meteorological concepts
Knowledge of the learning objectives	Synthesis and application
Learning and application	Understanding of how to
Objective knowledge	prepare for the FAA written
Opinion of the course value and	exam
instructor's ability to get the	Understanding of the subject
material across understandably	matter
Writing skills	

When asked about their academic evaluation based upon a mixture of Bloom's Taxonomy cognitive and affective domain key level words, the students responded with 54% "yes" responses to cognitive domain key words, as compared to 26% "yes" responses to affective domain key words. While this is considered a significant difference, and a key finding of this research study, there appears to be a lack of understanding among the students as to both cognitive and affective domain level meaning.

CONCLUSIONS

The researchers concluded that both the faculty and students surveyed were more familiar with the cognitive domain of Bloom's taxonomy than the affective domain. It was apparent that empowering the students by allowing them to choose their learning assessment tools and percentages had a powerful positive effect on the class environment and the learning outcomes.

Both the faculty and students chose to evaluate student learning through many varied techniques, in accordance with the literature review. It was apparent that both the faculty and students surveyed were satisfied with the student learning assessment policies at ERAU. The faculty and students differed on whether students were better evaluators of student academic learning than faculty. While faculty indicated that they assess more within the cognitive domain, students tended to indicate that they were assessed along both domains.

It was apparent that the faculty and students differed and had many opinions regarding what the central precept of students' learning assessment is, or should be, anchored to. Faculty and students alike chose the cognitive domain over the

affective domain by a wide margin, but their knowledge of the affective domain appears limited.

It was concluded that the survey instruments were invalid and unreliable for several areas of measurement and they should be revised extensively before further use.

RECOMMENDATIONS

The researchers recommend that ERAU provide faculty development to all Extended Campus faculty in student learning assessment, and that further research be conducted in this area, not only as a follow-on to this research study, but in expanded areas as well, including the use of Individual Evaluation Plans (IEPs). From the very positive effect noticed in the three classes which had student-chosen academic assessment, this technique should be studied further.

While most faculty and students surveyed were satisfied with their current ERAU academic assessment policies, the variety of assessment tools mentioned by both bears further study. It is recommended that the Extended Campus fund research on student learning assessment.

REFERENCES

- American Association for Higher Education. (2002). 9 principles of good practice for assessing student learning. Assessment Forum. Retrieved May 27, 2002 from <http://www.aahe.org/assessment/principi.htm>
- Anderson, J. (2001, March). Tailoring assessment to student learning styles. AAHE Bulletin. Retrieved on May 27, 2002 from <http://www.aahe.org/bulletin/styles.htm>
- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York: Longman, Inc.
- Bloom's taxonomy. (2002). Retrieved May 27, 2002 from <http://www.che.wsu.edu/~millerrc/bloom.html>
- Educational psychology interactive. (2002). Bloom et al's taxonomy of the cognitive domain. Retrieved May 27, 2002 from <http://www.valdosta.peachnet.edu/~whuitt/psy/702/cogsys/bloom.html>
- Innovative Teaching Concepts. (2002). Bloom's Taxonomy & multiple intelligence correlations. Retrieved May 27, 2002 from <http://www.twoteach.com/Bloom'sCorrelation.htm>
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). Taxonomy of educational objectives. Book 2: Affective domain. New York: Longman, Inc.
- Learning Skills Program. (2002). Bloom's Taxonomy. Retrieved May 27, 2002 from <http://www.coun.uvic.ca/learn/program/hndouts/bloom/html>
- Major categories in the taxonomy of educational objectives. (2002). Categories in the cognitive domain. Retrieved May 27, 2002 from <http://faculty.washington.edu/krumme/guides/bloom.html>
- Maki, P. L. (2002). Developing an assessment plan to learn about student learning. Retrieved May 27, 2002 from <http://www.aahe.org/Assessment/assessmentplan.htm>
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2001). Leverage points for improving educational assessment. CSE Technical Report 534. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing, UCLA.

SMHM 3500: Advanced field experience. (2002). Bloom's Taxonomy: Six cognitive levels of complexity in behavior: Verbs that help to describe complexity of behavior. Retrieved May 27, 2002 from www.scs.unt.edu/classes/smhm/3500/702/bloom.htm

National Center for Fair & Open Testing. (2002). Principles and indicators for student assessment systems. National Forum on Assessment. Retrieved May 27, 2002 from <http://www.fairtest.org/princind.htm>

Suskie, L. (2000, May). Fair assessment practices. AAHE Bulletin. Retrieved May 27, 2002 from <http://www.aahe.org/bulletin/may2.htm>

APPENDIX A

BIBLIOGRAPHY

American Psychological Association. (2001). Publication manual of the American Psychological Association (5th ed.). Washington, DC: Author.

Educational Testing and Measurement. (2002). Resource guide. Retrieved May 27, 2002 from

http://camellia.shc.edu/byrne/refguide/Testing_Assessment.htm

Louisiana State University. (2002). Centers for Excellence in Learning and Teaching (CELT). Retrieved May 27, 2002 from

<http://www.disd.lsu.edu/splash.htm>

Metropolitan Community College. (2002). Guide to theories of learning. Retrieved May 27, 2002 from <http://commhum.mccneb.edu/PHILOS/learntheo.htm>

Rosado, A, Dammier, E., Clark, R., & Rosenhammer, F. (Eds.). (2001). Guide to the graduate research project (5th ed.). Daytona Beach, FL: Embry- Riddle

Aeronautical University, Extended Campus.

APPENDIX B

SURVEYS

Faculty Coursework Evaluation Survey

Embry-Riddle professors Ron Clark and Jay Price are working on a research project that looks at the evaluation of college student performance in course work. Our work will be largely based on this survey, which will be randomly administered to ERAU faculty. Your assistance in completing this survey will provide invaluable, anonymous data pertinent to this research topic.

Thank you for your time and help. If you would like an executive summary of our findings, please provide your name and address below (your personal information will not be used nor reflected in our report):

Name: _____

Address: _____

Ronald Clark
Jay Price

Faculty Coursework Evaluation Survey

For items 1 through 9, either CIRCLE ONE OF THE ANSWERS provided or FILL IN THE BLANK.

1. Gender: Male Female

2. Age: _____

3. Non-teaching Occupation: _____

4. Courses you regularly teach at ERAU: _____

5. College Degrees held: _____

6. Number of years teaching for ERAU _____

7. Years of formal teaching experience: _____

8. Have you ever studied student evaluations? Hours: _____

**Cognitive and Affective Domain
Learning Assessment Choices**

9. Do you teach for other colleges or universities? Yes No

10. If yes, do the other colleges or universities have better student evaluation criteria or policies? Please comment: _____

11. Choose the type and value of evaluation method(s) you feel are best for the courses you teach (example: research paper: 50%; final exam: 40%; PPT: 10%)

- A. Individually authored research paper: % _____
- B. Jointly authored research paper % _____
- C. Case Study % _____
- D. Oral final exam % _____
- E. Take home open book final exam: % _____
- F. In class closed book midterm % _____
- G. In class open book midterm % _____
- H. PowerPoint presentation of paper % _____
- I. Verbal presentation of paper % _____
- J. _____ Quizzes in class % _____
- K. Other (specify) _____ % _____
- L. Other (specify) _____ % _____
- M. Other (specify) _____ % _____
- N. Other (specify) _____ % _____
- O. Other (specify) _____ % _____

For statements 12 through 18, CIRCLE A NUMBER from 1 to 7 that BEST DESCRIBES your opinion or experience.

Completely Agree 1 2 3 4 5 6 7 Completely Disagree

12. Under the current ERAU academic rules, I can accurately evaluate all of my ERAU students. 1 2 3 4 5 6 7

13. Students can evaluate each other better than faculty can. 1 2 3 4 5 6 7 _____

Completely Agree 1 2 3 4 5 6 7 Completely Disagree

14. My ERAU course grades have been based on my students' awareness and attention during class. 1 2 3 4 5 6 7

15. My ERAU course grades have been based on my students' responding to instruction in class. 1 2 3 4 5 6 7

16. My ERAU course grades have been based on my students' value choices during class. 1 2 3 4 5 6 7

Cognitive and Affective Domain
Learning Assessment Choices

Student Coursework Evaluation Survey

Embry-Riddle professors Ron Clark and Jay Price are working on a research project that looks at the evaluation of college student performance in course work. Our work will be largely based on this survey, which will be randomly administered to ERAU students. Your assistance in completing this survey will provide invaluable, anonymous data pertinent to this research topic.

Thank you for your time and help. If you would like an executive summary of our findings, please provide your name and address below (your personal information will not be used nor reflected in our report):

Name: _____

Address: _____

Ronald Clark
Jay Price

Student Coursework Evaluation Survey

For items 1 through 8, either CIRCLE ONE OF THE ANSWERS provided or FILL IN THE BLANK.

1. Gender: Male Female

2. Age: _____

3. Occupation: _____

4. ERAU degree program enrolled in: _____

5. College Degrees held: _____

6. Other colleges or universities attended: _____

7. Did other colleges or universities evaluate your academic performance better than ERAU currently does? Yes
No

8. If you answered question # 7 yes, how were you evaluated more favorably?

Please be very specific. _____

Cognitive and Affective Domain Learning Assessment Choices

Faculty responses to Survey Item #19

Ability to logically analyze problems and choose an appropriate solution method
Ability to think as a decision-maker
Comprehension of new material
Demonstrated ability to do the course work
Define, analyze, decide and present
Exams
Grasp of concepts and procedures
Knowledge of the course material
Knowledge of the learning objectives
Learning and application
Objective knowledge
Opinion of the course value and instructor's ability to get the material across
understandably
Perception of the presented material and application to their day-to-day
endeavors
Show an understanding of meteorological concepts
Synthesis and application
Understanding of how to prepare for the FAA written exam
Understanding of the subject matter

Faculty Responses to Survey Item # 21

Evaluating math and science is easy. At the undergraduate level, I am satisfied if the student can pick the appropriate methodology from those I present and apply it logically. This is about B+ level performance. I reserve an A for someone who that really requires some synthesis, just to see who can do it. I don't penalize someone who tries to apply the standard techniques to this problem, and consequently does not achieve a complete solution. This problem serves to "separate the men from the boys," if you will excuse the non-gender-neutral reference.

I perceive three general areas of difficulty when evaluating student's learning: 1. student personality and demeanor, 2. the impact of previous experience and learning, and 3. attendance vs effort.

I do not understand what you mean by the terms "value system", value complexes", "Value choices", and "valuing". Are these in Bloom's affective domain? If they are, can they be evaluated? How? The concept of andragogy is useful in the adult classroom. Students bring their own views and values and experiences to the classroom. Using andragogy, adults learn when they see a need. Using pedagogy, children are taught and are told what to learn.

Students need to learn how to evaluate, synthesize and apply information.

There has to be flexibility for individual instructors to evaluate students in a manner which is conducive to both the student and instructor. As widely varying as classes are, there are just as many methods for evaluating the student. I try to incorporate as many methods of evaluation as possible in order to capture as clear of a picture as possible of the student's knowledge level as well as their commitment to learning.

999 out of 1,000 surveys have "Strongly Disagree" to the left and "Strongly Agree" to the right. Terms in question 20 need to be better defined. What is a value system? Different things to different people.

In the courses I instruct, the end objective is not the same as that of traditional college courses. Conversely, the evaluation as to whether the end objective has been achieved or not, too must be in a form different from that which is traditionally utilized to evaluate the understanding of pertinent learning objectives. That is to say, the TRUE evaluation of success in the AMT program of study will be the results of the FAA written exams and the oral and practical exam given by the Designated Mechanic Examiner (DME), and issuance of an Airframe and Powerplant (A&P) Certificate. Therefore, my goal as the instructor/evaluator is to ensure that the students are grasping the knowledge required to overcome test anxiety, fear of public speaking and the ability to perform the practical projects which will be required of them by the DME. That is the basis of my evaluation process, to give the students the skills to help themselves pass the ultimate examination/evaluation. To date, the success has been quite good, only 1 failure out of 62 students to date (excludes current students and those who have not yet taken their FAA exams). It should be noted that the 1 failure did pass the exam on the next testing. Therefore, I believe that the current method of evaluation that I use is working quite well.

APPENDIX D

STUDENT DATA

Student Occupations

airline pilot
Industrial hygiene technician
aircraft mechanic
pilot/scheduler
USAF EWO
USAF
USAF Logistics
pilot
F-15 crew chief
flight engineer
pilot
USAF
USAF pilot
USAF
aircraft mechanic
USAF UAV pilot
security
pilot
shipping/receiving/ANG
student
operations agent
USAF
USAF pilot
airline captain
airport operations coordinator
maintenance officer
USAF weapons officer
sales manager
USAF fighter pilot
USAF officer
pilot
public safety officer

Student Indications of "Better" Student Evaluation at Other Colleges/Universities

It was just very specific numerical grades at SD (and a few other statistics). It gives you a better idea of exactly where you stand in relation to peers.

Cognitive and Affective Domain Learning Assessment Choices

A wider range of skills were tested with a higher workload. Also, evaluation was more frequent rather than having most of the evaluations come at the end of the class.

They were more concerned about giving academic credit where due. Not about making money.

Most classes were math class evaluation were cut and dry. For the type classes I've taken with ERAU I feel the evaluation process is favorable.

Student Responses to Survey Question # 17

application
attendance
communication skills (2)
development and value
GRP
knowledge (2)
meeting course objectives--learning the material
knowledge of course concepts
paper (4)
paper/briefs/test
paper with presentation (2)
participation (2)
presentation/research
research
research projects
responsiveness to the teacher's teaching
methods
tests and research papers
test scores (2)
the effort I put into each class
the quality of material I present or turn in to
class
work (2)
work completed
writing
writing skills

Student Responses to Survey Question # 19

I didn't understand what was meant by value choice, system or complex on previous page.

I do not know if the extended campus is different from the main campuses, but I would guess the courses are a little more relaxed. Otherwise I have enjoyed my time at ERAU-I just think that the grades come entirely too easy.

As this is my first course, I am not able to evaluate the grading process. However, I feel that I have learned quite a bit and will come out of this class knowing and understanding more.

ERAU has been great for my college education goals. I have time to do my job as an airline first officer and pursue my college education.

**Cognitive and Affective Domain
Learning Assessment Choices**

Being a student that has struggled in school and studies, I find that ERAU's intense subject matter is a great way to learn without distractions of learning useful information.

What is a value complex?

I feel that I have been evaluated by ERAU on how well I do presentations along with how well my final papers are. I feel this is an appropriate evaluation of how we (students) are to be judged. This is how the corporate world will be judging us.

