

RICHARD STOCKTON COLLEGE OF NEW JERSEY
EVIDENCE

PROGRAM ASSESSMENT FOR CONTINUOUS IMPROVEMENT

MARCH 2004

**IN THIS ISSUE of
EVIDENCE**

<u>Economics Program Assessment Plan</u> Deb Figart and Melaku Lakew	Page 1, 3
<u>Active (Direct) versus Passive (Indirect) Assessment of Learning Outcomes</u> Bonnie W. Buzza	Page 1, 4
<u>Performance Appraisals: Advantages and Disadvantages*</u> Gloria M. Rogers	Page 2
<u>Portfolios: Advantages and Disadvantages*</u> Gloria M. Rogers	Page 2, 5
<u>Assessment Woes</u> Mary Ann Trail	Page 5, 6
<u>How Can You Measure That?</u> Sonia V. Gonsalves	Page 6

* "Performance Appraisals" and
"Portfolios" were taken from "Assessment
Tools for Busy People"

**ECONOMICS PROGRAM
ASSESSMENT PLAN**

by Deb Figart and Melaku Lakew

Recognizing self-assessment as an integral dimension of the educational process, the Economics Program at Stockton College has implemented the following procedures and requirements. They are designed to further clarify the educational value of a Bachelor's of Arts degree in Economics and to gather information concerning the performance of economics majors over time (longitudinally).

Student Portfolio of Written Work

Each Economics major is required to submit one paper or project from an economics course each semester to be compiled in a portfolio of their work during the course of their study of economics at Stockton. Students may also choose to include other work—exams, other writing—in their portfolios. Upon completing the "Declaration of Major Form" declaring that economics is the chosen major or during the first semester of Junior Year, whichever is sooner, students will deliver their assembled portfolio-in-progress to the current Economics Coordinator. The portfolios will be maintained by the Economics faculty (the current Coordinator) and stored in a secure location. Beginning in their junior year, each economics major must make at least one submission to their portfolio during each semester at Stockton. The portfolio serves to illustrate the progress each student has made

(Continued on page 3)

Active (Direct) versus Passive (Indirect) Assessment of Learning Outcomes

By Bonnie W. Buzza

Associate Vice President for Academic Affairs; Member of Stockton's Assessment Task Force

When I wrote a previous article for *Evidence* I wrote about spending too much time gathering data without knowing in advance what you were seeking to measure. This present article suggests that you can learn some things from information serendipitously gathered or regularly gathered for purposes other than outcomes assessment, but it provides only a partial picture of whether or not a group is meeting its outcomes goals.

Passive, or indirect, assessment looks at data at hand, that seem to have some link to the outcomes you are seeking to understand, and draws conclusions. We at Stockton have a wealth of such data: for example, completion rates that are high, time

(Continued on page 4)

Performance Appraisals: Advantages and Disadvantages

Gloria M. Rogers

Definition: A *competency-based* method whereby abilities are measured in most direct, real-world approach. Systematic measurement to overt demonstration of acquired skills.

Advantages:

- Provide a more direct measure of what has been learned (presumably in the program).
- Go beyond paper-and-pencil tests and most other assessment methods in measuring skills.
- Preferable to most other methods in measuring the application and *generalization* of learning to specific settings, situations, etc.
- Particularly relevant to the goals and objectives of professional training programs and disciplines with well defined skill development.

Disadvantages:

- Ratings of student performance are typically more subjective than standardized tests.
- Requires considerable time and effort (especially *front-loading*), thus being costly.
- Sample of behavior observed or performance appraised may not be typical, especially because of the presence of observers.

Ways to Reduce Disadvantages

- Develop specific, *operational* (measurable) criteria for observing and appraising performance.
- Provide training for observers/appraisers.
- Conduct pilot-testing in which rate of agreement

(*inter-rater reliability*) between observers/appraisers is determined. Continue training and/or alter criteria for more specificity until acceptable consistency of measurement is obtained.

- Conduct observations/appraisals in the least intrusive manner possible (e.g., use of one-way observational mirrors, videotaping, etc.)
- Observe/appraise behavior in multiple situations and settings.
- Consider training and utilizing graduate students, upper level students, community volunteers, etc. as a means of reducing the cost and time demands on faculty.
- Cross-*validate* results with other measures, multiple methods should be used to *validate* the results of appraisals.

Bottom Lines:

Generally the most highly valued but costly form of student outcomes assessment. However, it is usually the most *valid* way to measure skill development.

Bibliographic References:

- Burke, Kay, ed. *Authentic Assessment: A Collection*. Illinois: Skylight Training and Publishing, Inc., 1992.
- Hart, Diane. *Authentic Assessment: A Handbook for Educators*. New York: Addison-Wesley, 1994.
- Ryan, Alan G. "Toward Authentic Assessment in Science via STS." *Bulletin of Science, Technology & Society*. 1994, v 14, n 5/6, p 290.
- Wiggins, Grant. "The Case for Authentic Assessment." *ERIC Digest*. December 1990.

Portfolios: Advantages and Disadvantages

Gloria M. Rogers

Definition: Collections of multiple student work samples usually compiled over time and rated using *rubrics*. The design of a portfolio is dependent upon how the scoring results are going to be used.

Advantages

- Can be used to view learning and development *longitudinally* (e.g., samples of student writing over time can be collected), which is a useful perspective.
- Multiple components of a curriculum can be measured (e.g., writing, critical thinking, research skills) at the same time.
- Samples in a portfolio are more likely than test results to reflect student ability when pre-planning, input from others, and similar opportunities common to most work settings are available (which increases *generalizability/external validity* of results).
- The process of reviewing and scoring portfolios provides an excellent opportunity for faculty exchange and development, discussion of curriculum goals and objectives, review of scoring criteria, and program feedback.
- Economical in terms of student time and effort, since no separate "assessment administration" time is required.
- Greater faculty control over interpretation and use of results.
- Results are more likely to be meaningful at all levels (i.e., the individual student, program, or institution) and can be used for diagnostic/prescriptive purposes as well.

(Continued on page 5)

Economics Plan

(Continued from page 1)

with respect to her or his economics skills: writing, analytical ability, problem solving, and critical thinking, for example. Further, these portfolios serve as a reference for faculty writing recommendations for students and for students who would like to make samples of their work available to prospective employers.

Senior Seminar

Each fall semester the Economics Program offers a senior seminar that all senior economics majors are required to take. Although the content of the course can vary from year to year depending on the professor, the basic structure of the course remains fundamentally consistent. Senior seminars in economics require students to:

- Give oral presentations on topics economic significance.
- Show evidence of facility using the electronic resources of the library, including but not limited to the sources of economics data available from the public and private sectors.
- Produce a 15-20 page Research Paper that properly references relevant primary and/or secondary literature on a topic. In terms of research methodology, both quantitative reasoning (use of tables, statistical tests, and/or regression analysis) and qualitative reasoning are required. Use of multiple regression analysis is strongly encouraged, though specific requirements may vary by semester or professor. (This paper *must* be submitted to the Student Portfolio of Written Work.)

Portfolio Rubrics:

4 Outstanding

Student paper with clear hypotheses
Logically coherent based on economic principles
Quantitatively sound
Assumptions clearly stated
Points backed by extensive references

3 Very good

Shows good understandings of the problem assigned but one of the elements listed in 4 contain minor errors

2 Satisfactory

Shows partial understanding of the problem assigned, but missing more than one elements in 4

Program Statement Regarding Student Learning

The Economics Program at Stockton College is dedicated to

creating an environment in which students can learn:

- To think analytically, critically, and reflectively.
- To communicate effectively orally and in writing.
- To read economic texts, articles, working papers, and arguments with an eye towards understanding, and to read critically, with an eye toward any economic ideology implicit in the writing.
- To reflect consciously and critically upon their own prejudices, opinions and beliefs.
- To present their own ideas in a clear and consistent manner and to be open to the ideas of others.
- To recognize the impact of the history of economic thought on the past and to bring the power of this thinking to bear on the future.
- To recognize the impact of globalization on the U.S and the economy of the developing world.
- To appreciate the power, fun and excitement of economic ideas and their very real impact on the world.
- To recognize the various economic policy and the tools available to economists to deal with various economic, social, political, urban, environmental and other problems.

Student Exit Interview Form

At the conclusion of the Senior Seminar or in their final semester at Stockton, whichever is later, Economics majors will be required to pick up from the Coordinator and complete the Economics Student Exit Interview Form that includes questions concerning each student's self-assessment of their familiarity with a variety of economics concepts. This form will be used to determine if students report a familiarity with the sorts of conceptual and methodological issues the faculty has designated as significant. It will also be used to gather information and feedback concerning the sorts of new courses students might be interested in having offered in the future.

The questions on the Student Exit Interview Form will serve as the core for a longitudinal study of our majors. The Economics faculty will devise a Graduate Self-Assessment Form that seeks to ascertain student feedback about their economics education and applications following graduation, e.g. 1-year out, 3-years out, 5-years out, and 10-years out.

Active versus Passive

(Continued from page 1)

to graduation figures that are low, and acceptances into grad schools and job placements that are substantial...some of which are anecdotal and unsystematized. We take this data to mean that we are succeeding here in our goals of preparing students well and efficiently for the next steps in their lives. This makes a sort of intuitive sense. Our students do better than others in the state on these measures. We conclude we are achieving the outcomes we have chosen (or been given by the state).

What we haven't factored in, of course, is that we have a selective admissions program that lets in primarily very good students with a solid chance of success at Stockton—and these students then succeed. And it has become more competitive to be admitted over the past several years. Had we an open admission policy and these same "output" (not "outcomes") figures, our passive data would be even more impressive! From this passive data we really don't know for sure if these students would have achieved equally well wherever they had gone to college, or—from this passive data—if Stockton has done anything to improve or expand upon what they knew about general studies or their major disciplines or the world at some distance when they arrived here.

However, we can feel good that our EOF students complete at rates equal to or above our regular admits—probably a tribute to the procedures and work of the EOF program, who have helped these students who come in at some risk and succeed here. Again, a use of passive or indirect data to draw a conclusion.

Considerable passive data is available at the college, and units can use it as a starting point to develop a more active assessment program. For example—we have enrollment figures over the years, and can generate class lists. What has been the demand for Alchemy 101? Has the demand changed—and if so, up or down? How about Alchemy 302? Have transfer students historically fared better in Alchemy 302 than our native-grown students? Have we let any students into 302 without the prerequisite—but who have had several Chemistry or Mystery of Life classes? How well have they done compared to the others? There are several sections of Alchemy 101 and of its continuation 102. Are there differences in the success rates of students in 102 based on the material they covered in 101? How about for Alchemy 201 based on 102? *[At a previous institution of mine the Modern Languages Department did such a passive assessment and discovered there were clear patterns—upon further inquiry they discovered little or no consistency among the 101 faculty for what they covered in their courses—thus students came into 102 unevenly prepared. Some then had several weeks of review before they encountered new material and others had gaps in their preparation. The department together established a set of learning outcomes or goals for all students taking 101, so that those who passed that course had*

a more consistent preparation for starting 102....and then they went on to establish learning outcomes goals between first and second years, and then second and third years, and for subsequent specialized courses.]

In contrast to passive assessment, active or direct assessment is intentional and is planned in advance. The program or unit determines what learning goals or outcomes it desires for its students, what processes should likely measure that learning, what degree of the learning it will accept as "success," and when and how exactly it will be measured (for example, each spring rather than once at the end of the students' careers here). Program A may feel that command of a certain vocabulary is critical for a student's truly having learned its discipline. The program may then create a vocabulary exam or use a standardized exam to test their students' knowledge. [As an aside, are there any patterns in the courses taken by those who succeed on the exam the first time, and those who don't? Does there seem to be a critical preparatory course that some students take and others don't? If so, should it therefore be required of all majors?]

Program B may take a more holistic approach. They may simply want their students to be able to apply what they've learned, whether or not they can describe it using a standard vocabulary. This program may use a senior project or laboratory exercise to evaluate a student's grasp of major themes or ideas from a rubric they've created for the discipline. [They might notice patterns in courses taken, as well.]

Whichever the decision of the program, in A or B there has been a discussion of the learning goals or outcomes desired and consideration of how best to determine how well they have been achieved. A plan has been put into place to measure them. The program then can base its curricular and pedagogical decisions on the information it has gained from this active or direct assessment. It has also used passive data to generate ideas for further assessment.

Both passive and active data have value. Passive data is often at hand and should be reviewed for conclusions it suggests. It can also suggest a direction for more active, planned assessment. There is always a leap using passive data for assessment, however. Intervening variables or alternate explanations may also account for the observed patterns.

Good programs make decisions based on available information, and active data gathered as part of a planned program of assessment, and thoughtfully analyzed, should help programs make even better decisions in order to improve student learning.

UPCOMING

**2004 AAHE Assessment Conference:
Connecting Public Audiences to Our Work
June 13-15, Denver
[Http://www.aahe.org/assessment/2004/](http://www.aahe.org/assessment/2004/)**

Portfolios

(Continued from page 2)

- Avoids or minimizes “test anxiety” and other “one shot” measurement problems.
- Increases “power” of maximum performance measures over more artificial or restrictive “speed” measures on test or in-class sample.
- Increases student participation (e.g., selection, revision, evaluation) in the assessment process.

Disadvantages

- Can be costly in terms of evaluator time and effort.
- Management of the collection and scoring process, including the establishment of reliable and valid scoring rubrics, is likely to be challenging.
- May not provide for *externality*.
- If samples to be included have been previously submitted for course grades, faculty may be concerned that a hidden agenda of the process is to validate their grading.
- Security concerns may arise as to whether submitted samples are the students’ own work, or adhere to other measurement criteria.

Ways to Reduce Disadvantages

- Consider having portfolios submitted as part of a course requirement, especially, a “capstone course” at the end of a program.
- Investigate the use of electronic portfolios as a means to increase process efficiency.
- Utilize portfolios from representative samples of students rather than having all students participate (this approach may save considerable time, effort, and expense but be problematic in other ways).
- Have more than one rater for each portfolio; establish *inter-rater reliability* through piloting designed

to fine-tune rating criteria.

- Provide training for raters.
- Recognize that portfolios in which samples are selected by the students likely represent their best work.
- Cross-validate portfolio products with more controlled student work samples (e.g., in-class tests and reports) for increased *validity* and security.

Bottom Lines:

Portfolios are a potentially valuable option adding important longitudinal and “qualitative” data, in a more natural way. Particular care must be taken to maintain *validity*. Especially good for multiple-objective assessment.

Bibliographic References:

- Barrett, H.C. (1994). *Technology-supported assessment portfolios*. “Computing Teacher,” 21(6), 9-12. (EJ 479 843)
- Hart, D. (1994). *Authentic assessment: a handbook for educators*. Menlo Park, CA: Addison-Wesley.
- Hodges, D. (1998). *Portfolio: A self-learning guide*. Barrington, IL.
- Paulson, L.F., Paulson, P.R., & Meyer, C. (1991). *What makes a portfolio a portfolio?* “Educational Leadership,” 48(5), 60-63. (EJ 421 352)
- Rogers, Gloria and Timothy Chow, “Electronic Portfolios and the Assessment of Student Learning,” *Assessment Update*, Jossey-Bass Publisher, January-February 2000, Vol. 12, No. 1, pp 4-6, 11.

Examples of an electronic portfolio system can be viewed at <http://www.rose-hulman.edu/ira/reps>

Assessment woes

Mary Ann Trail

After the Middle States visit we, the teaching librarians, were encouraged to begin assessment of our educational initiatives. We decided to start with a small, clearly understood project. We chose to evaluate the effectiveness of the Library's Research Skills workbook. This workbook is used in most Freshman Seminars to introduce freshman to the Stockton library.

We devised a 32 item, multiple-choice test to be used as a pre and post-test of the workbook. As a pilot project, we selected four Freshman Seminars to test the reliability of the test. Because we wrote a multiple choice test with many answers having more than five choices, we could not use the Scantron sheets to grade the test but had to grade it by hand. The data had to be entered by hand into the SAS program for analysis. We had a long wait for our results which proved to be more than a little disappointing. There was little significant improvement noted. Improvement seemed to depend on whether the student got one more question right after completing the workbook than before. It is difficult to demonstrate significant improvement if one is basing the change on such a small gain.

For the second year we changed the test. First we reduced the questions to 20, hoping students would feel less put upon. Then we worked very hard to find a way to make the test interactive using the WebCT quiz component. We thought students would be more interested in working on an electronic test than a paper one. We also wanted to do this for the ease of grading and faster access to the results. Also, in the future if the workbook is made interactive (as we hope) than the assessment component should also be electronic. For a variety of reasons the WebCT pre and post-testing proved to be a nightmare. As we found out, the soft-

(Continued on page 6)

Assessment Woes

(Continued from page 5)

ware is designed primarily to be used by students in a specific class. Entering a single class into the system is a simple process, quickly accomplished by the individual faculty member. Individual students must self-register using their STK number, not a quick and simple process.

Once we finally found four classes where the faculty members were willing to have us test their students, we had the usual problems of administering tests to freshman who don't really understand the need for assessment. This attitude was made clear in the analysis of the first two classes where not only was there no significant improvement but the results seemed to suggest the students were better off before the intervention. In one class, the grades were significantly lower after the workbook assignment.

Because we are convinced that we must have some measure of the effectiveness of instruction in library skills, we have resisted the temptation to say, "We tried it, it was too much trouble for too little results." We are still in the early stages of working out the kinks in the assessment of an important learning outcome. When we succeed, we will have a valid and reliable measure of what freshmen know about library skills before and after using the workbook. The struggle continues.

How Can You Measure That?

Sonia V. Gonsalves

Earlier this month Tim Haresign and I were co-presenters at a very well-received session at the Association of American Colleges and Universities (AACU) conference on Assessment in General Studies. As part of this presentation, we conducted an exercise to give participants practice in operationalizing affective outcomes in order to develop valid measures.

As part of the research for the presentation we looked at mission statements for several colleges' general education or general studies requirements and identified some common outcomes. Among these were stated goals for students to "appreciate diversity," "respect other cultures," and show "awareness of other cultures."

We examined the hierarchical levels of affective outcomes, receiving, responding, valuing, organizing, and characterizing by value as defined by Bloom, Kibler, and Gronlund, and discussed the usefulness of viewing these outcomes along a developmental continuum as goals for students.

The process of operationalizing any outcome starts with an analysis of the way the construct translates into a behavior. We asked participants to get together in groups and make a list of behaviors that would indicate an "appreciation of diversity." The exercise, in summary form, was to answer these questions:

A person who *appreciates diversity* does what?

- At work
- In class
- In the dorms
- With friends and family
- In spare time

A person who *appreciates diversity* –
Says what?

- Avoids what?
- Reads what?
- Goes where?
- Refuses to go where?
- Collaborates with whom?
- Shares with whom?
- Listens to what/whom?

These questions are to find out how "appreciating diversity" is expressed in a range of venues and in a variety of ways and are designed to get coverage of the expressions of this attitude.

The lists of responses and the questions themselves generated the information to begin the discussion of operationalizing the construct of "appreciating diversity." The end point that is desired is to have a representative list of observable outcomes that a group of informed people can agree are reliable measures of the construct. Some of these measures could be by self report, others by observation, and still others by test or performance. Some participants were heading back to their campuses to engage students in the process of operationalizing constructs.

If your program has affective outcomes that are posing a measurement problem, why not try operationalizing the constructs and see if that helps to clarify the measurement questions.

In the Next Issue of EVIDENCE

- **Assessing Information Literacy in Psychology**
- **Assessment in Student Services**
- **Student Engagement - A Second Look**