2006-2235: STUDENT SELF-ASSESSMENT: CAN IT BE USED TO IMPROVE INSTRUCTION?

Susan Miertschin, University of Houston

Susan L. Miertschin is an Associate Professor teaching in the Information Systems Technology program at University of Houston. Her teaching interests are in the development of information systems applications and the complementary nature of back-end developer and front-end developer skill sets. Her research is in the impact of instructional technology on student learning.

Carole Goodson, University of Houston

As a Professor of Technology at the University of Houston, Carole Goodson is the current chair of the assessment committee in the College of Technology. As an active member of ASEE, she is a member of the Academy of Fellows, a past editor of the Journal of Engineering Technology, a past Chair of the ERM Division , and a past Chair of the Gulf Southwest Section.

Luces Faulkenberry, University of Houston

Luces M. Faulkenberry is an Associate Professor and Coordinator of the Electrical Power Technology program at University of Houston. He earned a B.S. degree in Physics from University of Texas at Arlington and M.Ed. and Ph.D. in Industrial Education from Texas A&M University.

Barbara Stewart, University of Houston

Barbara L. Stewart is Professor and Coordinator of the Consumer Science and Merchandising, program at University of Houston. She earned a B.A. degree from Brigham Young University, a M.S. in Consumer and Home Economics Education from Utah State University, and an Ed.D. in Curriculum and Instruction from Brigham Young University.

Student Self-Assessment:
How Can It Be Used to Improve Instruction?

Introduction

A significant change to the culture of higher education is a broad-based, long-term focus on assessment of student achievement, course and teaching effectiveness, and overall program quality. The increased importance placed on assessment is evidenced by the recent US government appointment of a higher education commission charged with examining whether standardized testing should be expanded into universities and colleges to prove that students are learning and to facilitate comparisons of universities with respect to the quality of education they deliver.[1] While individual instructors have always had to be concerned with student assessment and processes for improving instruction, the scope of both assessment and instructor involvement with assessment has expanded. Instructors typically review not only students enrolled in their courses but also the courses they teach. However, large-scale assessments (e.g., program effectiveness) are more often the realm of administrators. Increasingly, instructors are asked to be involved with assessment of overall program quality, both for institutional accountability and for accreditation. Thus, it is important for freshman faculty to be familiar with a variety of assessment techniques as they begin their careers in higher education.

Student self-assessment is a technique that can be used together with other techniques to comprise an assessment effort. Student self-assessment refers to a student rating his/her own achievement of skills or knowledge. If new engineering educators encounter this technique as part of a program assessment approach, or if they wish to use it to help evaluate students or their class effectiveness, then an understanding of what it is, how it is developed, and why it is useful is imperative.

This paper examines factors that impact the effective use of student self-reports of learning achievement for improving curricula and programs. Areas to be addressed include the following.

- How can faculty effectively participate in a student self-assessment process? Are there advantages that accrue to the student when the educator uses this technique?
- What factors impact the validity of implementing this technique? Under what conditions and in what situations is it appropriate to use student self-assessment scores?
- An example of institutional use of student self-assessment is presented. The example includes a description of how an instrument was designed and how it is being administered. The development of the items is discussed as well as the development of the measurement scale. Limited data indicating how faculty expectations, as indicated by their responses to the instrument, compare to student responses. Attention is given to how student self-assessment results might be used to improve an instructional program.

Faculty participation in student self-assessment processes

A new engineering educator may be asked to participate in an evaluation process that includes student self-assessment for evaluation and validation of programs. The use of student self-reports of learning in this context is increasing as consideration is given to monetary and time costs of

implementing other types of assessment. Novice faculty members can contribute to a program assessment effort that includes student self-reports of achievement by becoming proactive in the process and knowledgeable about the measures applied. Among the ways that a faculty member can help are the following.

Participate in development of instruments. In consultation with peers, determine program priorities. Recognize that as self-assessment instruments are developed, program goals are defined that will subsequently apply to students and courses. Faculty are expected to achieve these goals and address the objectives. By participating in the process that determines the goals, new engineering educators have the opportunity to include areas that they believe are valuable to students.

Monitor the quality of instruments and the administration process. If an instrument is to be used to make judgments regarding class content and faculty performance in content delivery, it is important to maintain the quality of the instrument and the process. Periodic review of the instrument is required to assure it remains relevant and continues to measure what is intended. Be involved with the review process. Look for ways that the administration process might bias the results and make these concerns known.

Participate in policy formulation and review with respect to results. Be proactive in formulating policies regarding the way results of evaluations are used and distributed. Become active in departmental assessment bodies or embrace other opportunities for faculty involvement. Participate with colleagues in discussions concerning use and interpretation of assessment data.

Help ensure the integrity of the process. Any single assessment tool or process should not be used in isolation. Multiple measures should be used and cross-validated, rather than relying on a single instrument. Evaluate the overall assessment process to ensure that it is appropriate and that stakeholders have a voice when courses and programs change. Remember that students are also important stakeholders and that, often, students have a voice only through faculty advocates.

Try different course assessment techniques. An excellent way of becoming more familiar with different assessment techniques is to try them with respect to student achievement in one's own classes. Use methods other than tests, or use tests to assess the effectiveness of different teaching techniques. A faculty member might compare student performance on tests from year to year – after normalizing or statistically controlling for differences in the initial academic backgrounds of the classes being compared. Student backgrounds at the beginning of class might be measured using a student self-assessment technique.

Use the results. Too often results of assessment are set aside with little thought or reflection until an accreditation team or university review team requests the data. Instead, carefully consider the results in a timely manner. If possible, use the results to identify program strengths and weaknesses. Then make improvements in programs and/or classroom instruction. In reviewing results, consider whether the instrument itself needs modification.

In parallel to the faculty member becoming more familiar with assessment through personal involvement, an effort could be launched to aid the students in becoming more accurate in

assessing their own academic growth and achievement. Steps to this end might include supporting on-going experience with self-assessment in classes, perhaps as an aid in guiding students with their study efforts.[2] Research supports the importance for students to develop a capacity for reflective assessment of their own activities, explaining that self-reflection helps with development of understanding.

Factors related to validity with respect to self-assessment instruments and/or processes

There is a large body of literature that addresses educational uses of self-assessment. This literature examines self-assessment in two contexts; one is its use with individual students for one of two purposes. It is used either to enhance learning through self-reflection or to provide a summative assessment for evaluation of the student's achievement. A second context for use of self-assessment is for evaluation and validation of programs. Within either of these two contexts, an educator wants to know whether self-assessment is valid; i.e., does the technique accurately measure student achievement?

Student self-assessment of achievement is attractive for its relatively low cost and ease of administration. However, it is useful only if it is reliable and valid. In a summary of literature related to student self-reports of learning Volkwein notes that self-assessment measures, in general, have only moderately positive correlation with objective measures when used to gauge the learning or skill of individuals. However, he reports that results of these same methods have fairly high reliability when aggregated and used to compare the performance of groups.[3]

Although the question of validity of measures is of general concern, the self-assessment approach is particularly controversial and, under scrutiny, has produced divergent results. For example, a study by Sarin and Headley reported a relationship between test scores and corresponding self-assessments for a group of industrial engineering students.[4] A more recent summary work by Sundström found thirty-three out of forty-three studies that demonstrated that self-assessments and objective test scores or other external rater evaluations were equally competent in predicting academic performance.[5] That leaves roughly one-fourth of the studies reviewed showing discrepancies between self-assessment results when compared to the results of objective measures. Thus, while results vary, there is evidence that students can rate themselves with respect to achievement with competence comparable to that of an expert (such as a teacher) under some circumstances. The large body of literature that addresses the conditions that impact the educational uses of self-assessment indicates that self-assessment is more valid under certain conditions. A research summary categorized by factor follows.

Results of self-assessment can depend on the ability of the student. In a critical review of fifty-five studies related to self-evaluation, Mabe and West found support that more intelligent and capable students have a better perception of their ability-related performance and make more accurate judgments concerning the level of their performance.[6] Boud and Falchikov concluded from their review of more than sixty studies that there does seem to be a tendency for high-achieving students to be more realistic in self-assessments, thus sometimes underestimating their performance, while low achievers tend to overestimate their performance.[7] The recent study by Sarin and Headley confirmed this finding in their study of high-performing versus low-

performing industrial engineering students.[4] Thus, it appears that higher-achieving students more accurately assess themselves with respect to achievement.

Results of self-assessment can depend on the class level of the student. Boud and Falchikov, from their critical review of more than sixty studies related to self-assessment in higher education, found that upperclassmen and graduate students tended to become either more accurate in rating themselves or tended to underestimate their performances in comparison to freshman or sophomore undergraduates.[7] This finding complements the finding that higher-ability students are able to more accurately assess their achievement, since it is likely that more higher-ability students are successful in attaining the rank of upper-classman or graduate student.

Results of self-assessment can depend on student experience with the technique. One factor considered in previous work is practice. That is, does a student's ability to accurately self-assess improve with practice? Maskell cites a study supporting improvement over time. He found that there was little relationship between student ratings and instructor scores among first year students in electrical engineering, followed by a strong relationship between these same variables for the same students in their second year.[8]

Results of self-assessment can depend on the nature of the skill being evaluated. Coombe points out that the degree to which learners are capable of valid self-assessments likely depends on the nature of the skills being assessed. Concepts and skills dependent on a complex learning process (such as language acquisition) are difficult to self-assess; furthermore, the accuracy of self-assessments will depend on the accuracy with which either learners or experts can define the skills they are to assess in concrete behavioral terms and develop effective assessment tools.[9]

Results of self-assessment can depend on the measurement scale. Sundström summarizes research indicating that self-assessments are affected by the scale used to measure them. People tend to rate themselves above average when they use relative scales, but when an absolute scale is used and when the questions are specific, rather than general, the self-assessments are more realistic.[5]

Recommendations exist for conditions that support using student self reports as substitutes for achievement test scores. The National Center for Higher Education Management Systems (NCHEMS) provides four criteria for evaluating the use of self reports of academic development as proxies for achievement test scores: (1) the measures should represent broad-based outcomes, (2) the measures should represent significant phenomena that can be used to inform policy actions; (3) the measures should reliably covary with other assessments; and (4) the observed relationships should persist across different educational settings.[10]

The recommendation of criteria to guide the use of self reports of academic development as proxies for achievement test scores inspired research by Pike. This research was designed to empirically verify the recommendations as sufficient conditions for using self reports in place of achievement test scores. Pike compared student responses to self-assessment items with scores on the College Basic Academic Subjects Examination (College BASE). The study established support that student self reports could be used as proxies of achievement tests, but also raised issues that Pike explored in further in a subsequent study.[11]

In the subsequent study, Pike advanced and explored two reasons why some studies show low correlations between student self reports and measures of academic achievement. First he reported that self-reports tend to measure generic college outcomes while achievement tests measure more specific skills. Second, he established that objective achievement tests generally are designed to accurately measure a relatively narrow range of behaviors, while self reports tend to measure a broader range of behaviors with relatively low precision; these measurement method differences give rise to method-specific score variance that attenuates correlations between self reports and test scores. Finally, Pike made the case that the first difficulty is remediable through careful matching of the instruments to content, while the second difficulty cannot be remedied by instrument design; it is a fundamental difference in two measurement techniques. The research also led him to conclude that in order for achievement test scores to serve as proxies for self-reports, these self-reports must measure the same constructs as the achievement tests. Specifically, his research showed that self reports using an instrument based on the same specifications as College BASE was justifiable when used as a general indicator of achievement for two-year and four-year colleges and universities, but he cautioned against extrapolating the validity to other instruments and/or across other types of institutions.[12]

An example of institutional use

In order to assess its programs for curriculum improvement, the College of Technology at the University of Houston compiles and analyzes student self-reports of concept and skill achievement just prior to graduation. This has been accomplished largely by the Assessment and Continuous Improvement Committee (ACI), which was formed to plan and implement program assessment for diverse program areas within the College. The ACI developed the Graduating Senior Survey (GSS) instrument in-house with help from professionals in the university's Learning Assessment Services. The GSS instrument measures student perceptions of their own learning in two areas: (1) program specific content and (2) broader educational goal areas. The GSS items specific to particular programs were developed by the faculty from the specific program area. Although the GSS is an indirect measure of student performance that relies on student self-assessment, experience to date has demonstrated that the results are informational with respect to focus areas in need of improvement.

The instrument that was used and reported on herein includes three sections. Section 1 asks for demographic and employment information. Section 2 addresses competence in curriculum areas that are required of all students in the college including computer usage, written communication, oral communication, analytical skills applied to content area, project management, continued learning, and team functioning. The items in Section 3 relate to competencies within specific programs. All items were either refined from previous baseline items or newly developed by the ACI with input from program faculty. Early administrations of the survey indicated that a Likert scale did not produce adequate discrimination in response, so the committee researched other scales. The response scale used was adapted from a classroom checklist proposed by Angelo and Cross.[13] The students rate their acquisition of a stated skill or concept area as A Advanced level of skill and knowledge, S Satisfactory/Functional skills and knowledge, B Basic skills and knowledge, or N No skills or knowledge. In fact, the new scale is more successful in producing differentiated responses. Administering the survey each spring, with the assistance of the faculty in senior level courses, ensures a good return rate.

The ACI reviewed the instrument of this example relative to the criteria proposed by the NCHEMS. They determined that it satisfied at least two of the criteria that support use of self reports of academic development. The outcomes measured by the instrument are broad based outcomes and the measures represent phenomena that can be used to inform policy actions. The use of broad based instruments is consistent with Pike's finding which indicates that self reports are justifiable when used as a general indicator of achievement in colleges and universities. In order to further justify the use of the instrument, the ACI plans to examine the relationship with other objective student assessment measures in the future.

What are faculty expectations of student responses? Table 1 and Figure 1 show the expectations of faculty compared to student perceptions of knowledge for each of the items that generate college-wide data (as opposed to program-specific data). From the table and graph, several general observations can be made.

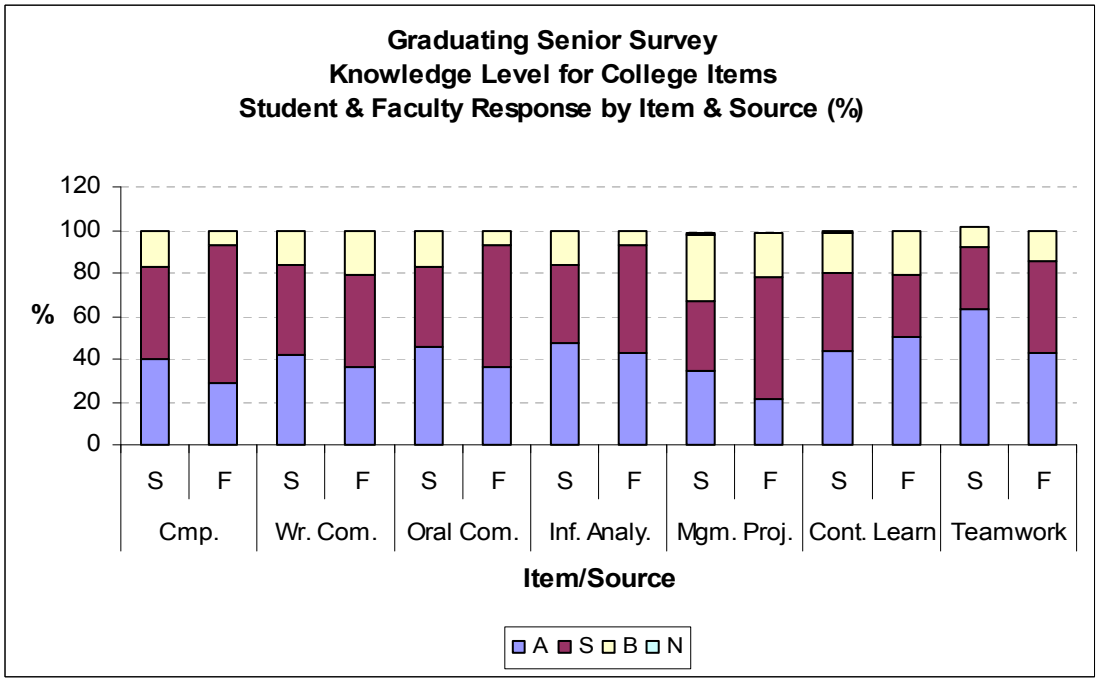
- In general, the students' perception of their knowledge base is consistent with the faculty members' expectation.
- Areas with the highest faculty expectation involved computer proficiency, oral communication, plus the analysis and interpretation of information in the student's major field of study. Computer proficiency reflected one of the greatest discrepancies between faculty expectations and student perception, with faculty expecting a higher level of proficiency than students perceive.
- The area reflecting the lowest faculty expectation involved the management of projects. This finding is consistent with student perception in that this area reflected the lowest student perception. However, this area also reflected one of the greatest discrepancies between student perception of knowledge and faculty expectation.
- The area with the highest student perception involved the ability to work on teams. The faculty have somewhat lower expectations than student perceptions.

Table 1
Knowledge Level for College Items
Student & Faculty Response by Item (%)

Item	Source	A	S	B
Computer Proficiency	S	40	43	17
	F	29	64	7
Written Communication	S	42	42	16
	F	36	43	21
Oral Communication	S	46	37	17
	F	36	57	7
Information Analysis	S	47	37	16
	F	43	50	7
Manage Projects	S	34	33	31
	F	21	57	21
Continue Learning	S	44	36	19
	F	50	29	21
Teamwork	S	63	29	9
	F	43	43	14

A: Advanced, S: Satisfactory, B:Basic

Figure 1



How can student self-assessments be used to improve the instructional program? In general, the ACI believes the GSS instrument has provided some valuable general feedback regarding elements that need review and enhancement. The instrument has been found to be convenient to use and flexible in format.

The ACI members agree that when the achievement expectations of the faculty in a specific area are greater than the student self perceptions reported on the GSS, additional investigation is merited. Program faculty members are the proper people to examine divergences in items related to a specific major. In reviewing survey results, the faculty must ultimately decide (1) if the results warrant curricula revision, (2) if a survey item needs revision, (3) if faculty expectations are appropriate, and/or (4) if the instrument is fairly measuring the concept. The program faculty, both new and veteran, must bring their experience and expertise to the investigation for the answer. When self-reported student achievement is greater than the faculty expectations, it must be asked if we are indeed doing a superb job; that is, do the students really exceed our expectations or should we further evaluate the measurement?

The best information will be achieved as the responses are tracked over time. The data collected will provide information regarding the general direction of the programs and alert faculty to both long-term program progress and short-term problems.

References

1 Arenson, Karen W. Panel explores standard tests for college. *The New York Times*. Feb 9 2006.

2 Curtz, Thad. Teaching self-assessment. Last accessed at: http://www.evergreen.edu/washcenter/resources/acl/e1.html on March 1 2006.

3 Volkwein, J. Fredricks. On the correspondence between objective and self-reported measures of student learning outcomes. 2005. Last accessed at: http://www.ed.psu.edu/cshe/abet/pdf/Self_Reported.pdf on December 1 2005.

4 Sarin, S. and Headley D. Validity of student self-assessments. *Proceedings of the 2002 American Society for Engineering Education Annual Conference and Exposition.* 2002.

5 Sundström, Anna. Self-assessment of knowledge and abilities: a literature study. *Educational Measurement, 54.* 2005.

6 Mabe III, P. A. and West, S. G. Validity of self-evaluation of ability: a review and meta-analysis. *Journal of Applied Psychology, 67, 3.* 1982. 280-296.

7 Boud, D. and Falchikov, N. Quantitative studies of student self-assessment in higher education: a critical analysis of findings. *Higher Education, 18.* 1989. 529-549.

8 Maskell, Doug. Student-based assessment in a multi-disciplinary problem-based learning environment. *Journal of Engineering Education*. April 1999. 237-241.

9 Coombe, Christine. Self-assessment in language testing: reliability and validity issues. January 2002. Last accessed at: http://www3.telus.net/linguisticissues/selfassess.html on November 14, 2005.

10 Ewell, P. T., Lovell, C. D., Dressler, P., and Jones, D. P. (1994). A preliminary study of the feasibility and utility for national policy of instructional "good practice" indicators in undergraduate education. Washington D. C.: U.S. Government Printing Office. NCES 94-437.

11 Pike, Gary R. (1995). The relationship between self-reports of college experiences and achievement test scores. Research in Higher Education, 36, 1-22.

12 Pike, Gary R. Limitations of using students' self reports of academic development as proxies for traditional achievement measures. *Annual Forum of the Association for Institutional Research Paper.* May 1995.

13.Angelo, T. A. and K. P. Cross. *Classroom Assessment Techniques: A Handbook for College Teachers*, Josey Bass, San Francisco: 1993, 285.