View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by AIS Electronic Library (AISeL)

Using an Assessment Exam to Assess a CIS Program

Completed Research Paper

Diane Lending
James Madison University
lendindc@jmu.edu

Abstract

Assessment of academic programs is required by accrediting bodies, universities, and colleges. The objective of this paper is to share one method of assessing a CIS program; an in-house developed assessment exam. The paper describes the development, reliability, validity, and use of an assessment exam used in a CIS program at a public university in the Mid-Atlantic region of the United States. The exam is given to all students pre- and post-program. The faculty within the program developed the assessment exam to correspond to the objectives of the program. Developing the exam required a great deal of time by the faculty members. The discussions in developing and maintaining the exam have helped to strengthen the program and the bonds between faculty members. The results of the exam have been used to improve the courses and the whole curriculum.

Keywords (Required)

Program Assessment, Assessment Exam, IS Program

Introduction

Assessment of academic programs is required by universities, accrediting bodies, and other constituents. The focus of assessment should be on making continuous improvements to the program (Merhout et al., 2008.) Doing program assessment well consists of developing objectives, assessing these objectives, evaluation of assessment results, and, most importantly, improvement of the program based upon the results. Assessment techniques can include direct and indirect measures, quantitative, and qualitative measures, and course-embedded and overall program measures. The course-embedded techniques are a natural part of a course, perhaps part of an exam or project, which is re-purposed to be used to evaluate the course. Program measures can include indirect measures such as surveys or focus groups or direct measures such as exams. The best assessments include multiple complementary techniques (Jacobson et al., 2010.)

Program assessment exams can be purchased or developed internally by a program. The advantage of a standardized purchased exam is that it provides an independent validation of the program, it can be implemented quickly without substantial faculty effort, and it provides benchmark data to compare to other institutions (Jacobson et al., 2010.) On the other hand, the objectives assessed in a purchased exam may not match the objectives of a program, the cost of such an exam can be prohibitive, and the lack of faculty effort may lead to less buy-in by the faculty to accept any results (Jacobson et al., 2010.)

This paper describes an assessment exam used in a CIS program at a public university in the Mid-Atlantic region of the United States. The exam is a part of multi-faceted assessment program which is summarized in Table 1. All assessments except the alumni survey are conducted annually.

Assessment type	Assessment	Level
Direct Measures	Course-embedded assessments in each required course	Course
	Assessment exam	Program
Indirect Measures	Focus group of graduating seniors	Program
	Survey of graduating seniors	Program
	Alumni survey	Program

Table 1. Summary of Assessment Methods

Exam Development

The program assessment exam was developed by the program faculty. We considered using a standardized assessment exam. However, the lack of correspondence with our program objectives was seen as a major issue. We conducted a sample course assessment using a standardized exam and the results were unsatisfactory. Students were openly frustrated with an exam that included questions on issues that were not taught in the course. The instructor piloting the exam felt that the result could not be trusted as the course objectives and the exam objectives did not have sufficient overlap. Thus the exam was developed internally by the CIS program faculty to directly correspond to the program objectives and course objectives. There are multiple questions per major program objective.

The exam was first written in 2004 and has been given annually since then. Minor revisions are made upon seeing results. A major review of all questions is conducted every two or three years. The current version of the exam has 60 questions.

For each question, we categorized the question as either problem solving or terminology. We show the major learning objective and sub-objective. We explain why each of the distracters is wrong. A sample question, for Systems Analysis and Design, is included in Table 2. This description is done for each question and has been reviewed by all faculty members who teach courses that meet this objective as well as other interested faculty. The faculty members review the items and agree that the items matched the objectives as intended.

The amount of work involved in developing the exam was considerable; however, the faculty members concerned feel that the effort has been invaluable. In a later section of this paper, we discuss the value of the results, but there is great value simply in the discussions that occur in developing and evaluating the exam questions. For example, development of questions by faculty teaching the same class, leads to important discussions on what each is teaching, how, and when. Faculty teaching classes that form a sequence discuss what is being taught in the earlier class and how it influences what is being taught in later classes. These discussions lead to changes in curriculum, objectives, and emphasis. Just developing the exam helps bring about a program that consists of faculty working together to develop a cohesive curriculum. The periodic review of exam results and questions keeps the discussions continuing.

Davis, a systems analyst, needs to know detailed information about the sales management process. He has been told not to get involved with the integration of sales information with other departments, just to collect rich information on the sales management process. He needs to get manager's opinions of the current process that will assist him in designing features for a new system. The appropriate analysis technique to be used is _____. a) interview b) document analysis c) observation d) questionnaire
Question Type: Problem Solving
Learning Objectives: Analysis: Select the appropriate requirements gathering technique for a system development project.
Explanation of distracters: a. correct b. document analysis is not an appropriate requirements-gathering technique for gathering rich information. c. observation does not provide for rich information gathering and is not an appropriate requirements-gathering technique for designing features for new or to-be systems. d. questionnaires are not an appropriate requirements-gathering technique for rich information

Table 2. Sample Assessment Exam Question

Pre-Program Post-Program Model

We recently moved to a pre-program, post-program model of using the assessment exam. The pre-program exam was given as a quiz in an early class in the program. The post-program exam was given as an exam in our capstone class. All students taking the classes take the exam. Both courses are required courses in the major so every student must take them. This allows us to assess whether the program is having the desired effect on learning.

The students received a grade in the classes for the exam. The pre-program grade was curved significantly since we do not expect students to do well on the exam since they had not taken the classes where they would learn the material.

Notice that this is not a true pre/posttest. The students taking the pre-program exam are not the same students taking the post-program exam. Recent research (Yorke & Zaitseva, 2013) suggests that this method of assessment is an adequate substitute for longitudinal data. Over time, we will be able to follow the same student pre- and post-program but until this upcoming year, it was not possible. In prior methods of test administration, we found that a student who took the exam at different times in their program scored substantially higher on the second exam.

The students in our program can declare their major at any time during their college career; however, based upon their academic record, they are officially accepted into the major at the completion of their sophomore year. Assuming a typical progression through the program, the students who take the pre-program exam are beginning of their junior year. They have usually taken an introductory MIS course and

a Visual Basic programming class before this class. The students taking the post-program exam are in the last week of their last required class in the major; often, immediately before graduation.

Data Analysis

Reliability and Item Analysis Calculations

We begin with a reliability calculation where we calculate Cronbach's alpha for all 60 questions on the Assessment exam. This past year Cronbach's alpha was .77 (N=137) indicating a reliable exam.

Item analysis included a correlation with the total scores. This correlation is done separately for the pre- and post-program results since the pre-program results may be little more than guessing. Item difficulty (or "easiness") ranges from 0 to 1 and represents the proportion of students who answered the item correctly. Difficulty values closer to 1 indicate the item was easy; difficulty values closer to 0 indicate the item was difficult. An item that is negatively correlated with the total score raises a red flag. The faculty members who teach the classes in question discuss that item. Reasons found recently are that the material is no longer being taught, the terminology has changed, or the question is too difficult. Again, simply discussing the items leads to greater cohesion of curriculum.

We calculate reliability of questions for each subscale (question related to an objective.) For each of the subscales, the Cronbach's alpha varied from .34 to .61 which is relatively low. However, Cronbach's alpha tends to be low for questions that are scored as either right or wrong. For formative purposes, this assists us in investigating whether any of the subscale questions need to be revised.

Expected Results

We expect that students who are post-program will do better than students who are pre-program. We expect students who have taken a class to do better on subscale questions than student who have not taken the class. While these expectations may seem obvious, simply investigating whether this is true has been shown to have great value. In addition, we have found cases where this is not true and that has led to changes in curriculum and exam.

Results from Exam administered in 2013

As expected, pre-program scores average lower than post-program scores as shown in Table 3. Cohen's *d* calculates as 2.42 showing a huge effect of taking courses within the CIS major on these scores.

Objective	Results			
	Pre-program n = 98		Post-program n = 42	
	Mean correct	SD	Mean correct	SD
Overall Score	43.52	8.56	62.94	6.85
Programming	45.78	14.73	64.13	11.11
Database	41.75	15.76	57.74	10.64
Systems Analysis	45.61	17.41	80.48	10.11
Architecture	49.32	15.03	65.08	13.55
Telecommunications	34.14	15.39	48.70	14.73

Table 3: Comparison of Pre- and Post-Program Results

Our interpretation of these results is that completing the CIS program has a strong effect on scores on the assessment day exam. There is little overlap between scores pre- and post-program.

In Table 4, we used the course(s) that mapped as having substantial coverage of the objective being assessed and compared the score on that subscale for students that had and had not taken the course. As expected, subscale scores were significantly higher for students who had taken the relevant course(s) than for those that had not taken the relevant course. Cohen's *d* suggests that the "treatment" of taking the course, substantially changed students' performance on these questions. We conclude that each course is teaching the material that we believe it should be teaching and covering the objectives that we designed it to cover. In telecommunications, the post-program score is below 50%. This result led to a very productive discussion among the faculty who taught the course as to what material taught in the course should be retained a year after the student took the course.

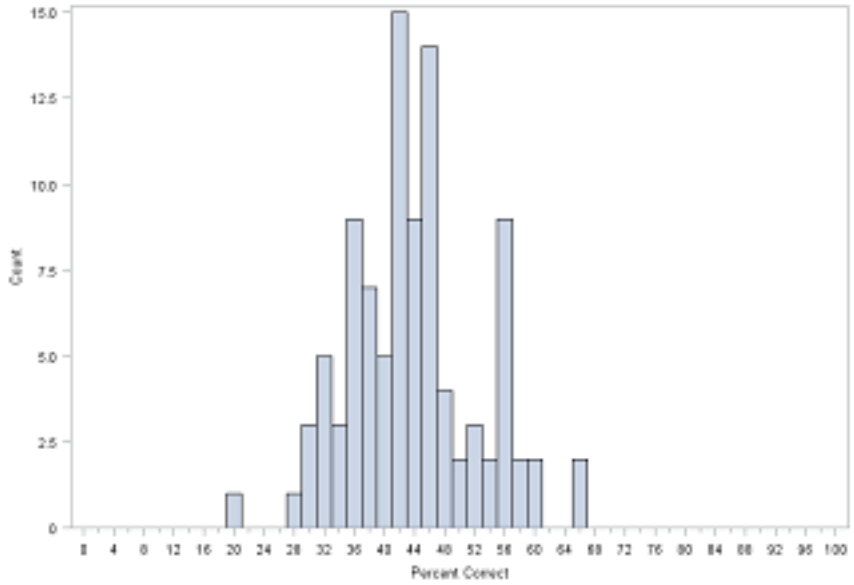
Further analysis of the pre- and post-program scores is shown in the two graphs shown in Figure 1. The first shows the distribution of percent correct scores for the pre-program scores. Note that the vast majority of students pre-program score below 50%. The second shows the distribution of percent correct scores for the post-program exams. Note that the vast majority of students post-program score above 50%.

One area of concern is that even the post-program scores are not particularly high. The highest score this year was 78. The program faculty has begun a discussion about what students can be expected to remember a year after a course has completed.

Objective	**Results**						**Taking course had an effect on score?**		
	Have Not Taken Course			Have Taken Course			Pooled t-value and p value	Cohen's *d*	Interpretation
Program-ming	N	Mean correct	SD	N	Mean correct	SD			
Course 1	40	42.33	14.34	92	56.23	14.84	t = 4.99 p < .0001	*d* = .95	Statistically significant difference. Large effect
Course 2	73	43.01	13.01	41	64.07	11.24	t = 8.69 p < .0001	*d* = 1.71	Statistically significant difference. Huge effect
Course 3	99	45.99	14.80	41	64.07	11.24	t = 7.02 p < .0001	*d* = 1.31	Statistically significant difference. Very large effect
Database	63	40.08	16.04	44	57.58	10.76	t = 6.31 p < .0001	*d* = 1.25	Statistically significant difference. Very large effect
Systems Analysis	95	44.10	15.40	15	80	12.54	t = 8.58 p < .0001	*d* = 2.41	Statistically significant difference. Huge effect
Architecture	32	50.52	14.81	42	65.08	13.55	t = 4.40 p < .0001	*d* = 1.05	Statistically significant difference. Large effect
Telecomm-unications	88	32.64	15.14	26	49.30	15.68	t = 4.89 p < .0001	*d* = 1.1	Statistically significant difference. Very large effect

Table 4: Analysis of Objectives Based Upon Courses Taken

Pre-Program Scores



Post-Program Scores

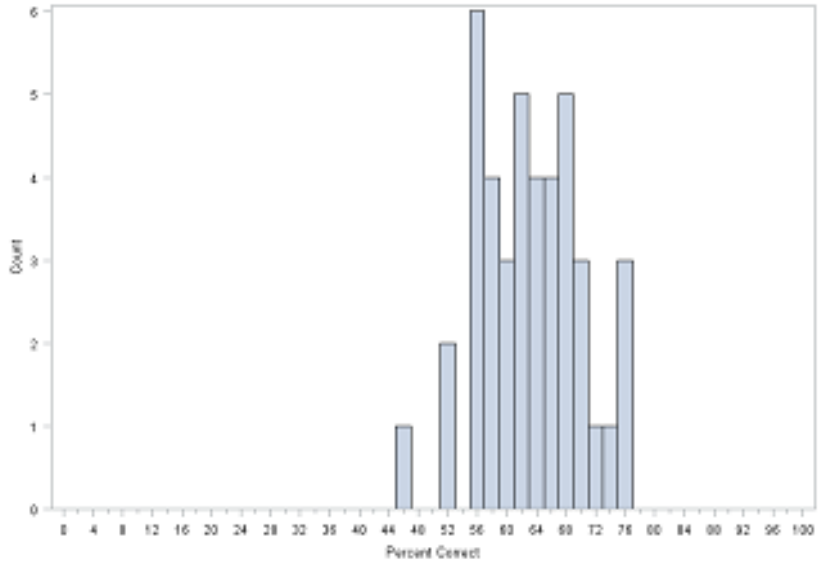


Figure 1: Comparison of Score Distributions

Validity of the Assessment Exam

There was a moderate relationships between total scores and cumulative GPA ($r = .32$, $p < .001$, $N = 107$).

The graphs of correct scores pre- and post-program also suggest that the exam is valid. The students who have just begun the curriculum receive substantially lower scores on the exam than those who are about to complete the curriculum.

Discussion

Using Results

The greatest value of assessment comes when the results are used (Jacobson, et al., 2011.) In our program, the CIS faculty members involved in each course or course-sequence discuss the assessment exam results annually. As previously mentioned, we have seen a great value in just having the discussions. However, the important next step is to use what is learned. Once faculty members understand the point of the assessment process, it is possible to use the results in changing curriculum because of the assessment exam. For example in one particularly illuminating administration of the exam, we saw scores on the design portion of the exam lower post-program than pre-program. That change coincided with a change in Systems Analysis and Design textbook. After much faculty discussion, we ended up writing a supplement to the book to be used in all sections of the course. The exam scores bounced back after that intervention.

Motivating Faculty

An assessment exam such as ours requires a great deal of faculty effort. Thus it is critical to have faculty buy-in to the process (Jacobson, et al., 2011, Merhout, et al., 2008.)One way to motivate faculty is through encouragement by the department, college, and university. Another motivating factor is accreditation which requires assessment. However, the results themselves are an important motivating factor for our faculty. Once the process becomes routine, assessment can become self-motivating. Being able to use the results and having the discussion can become an important motivating factor in-and-of itself. We feel that having our own assessment exam rather than a standardized one aids in this since the questions are tied to what the faculty members expect students to learn in their courses.

A second critical piece in making this assessment work is having a strong assessment leader (Jacobson, et al., 2011, Merhout, et al., 2008.) Since our assessment exam has been developed, we have had two different assessment chairs, both senior members of the CIS faculty. The chair must have the persuasive skills to keep all faculty members on board with the process and the organization skills to keep the testing on track. This task does require a substantial effort by the assessment chair but our institution has recognized and rewarded the effort at the department, college and university level. Additionally because our institution rewards pedagogical research, both chairs have been able to write about and use their efforts in research.

Motivating Students

Another important piece of making this assessment process work is to make sure students buy in to the process. If students are not taking the assessment exam seriously, the results simply will not be valid. Our program has gone through several methods of exam administration to settle upon one that works. One method of student motivation is to make the exam a graded part of the class. This, especially in the post-program assessment, leads to students taking the exam seriously. It is more difficult to achieve this pre-program since the students will miss many questions. We have handled this via having the faculty member who administer the exams explain why the testing is done. The story about the design supplement to our systems analysis and design class has proved to help with student motivation. When students understand that we are truly interested in improving the program and not in evaluating them, they take the exam much more seriously.

Conclusion

This paper reports on the successful assessment exam developed and administered in our CIS program. We feel that the major effort involved developing and maintaining the exam has paid off with increased communication within the department. The results of the exam and the discussions about the results have been used to strengthen the program and the bonds between faculty members. The results of the exam have been used to improve the courses and the whole curriculum.

Acknowledgements

The author would like to thank Mike Mitri for his leadership in the initial development of the assessment exam and George Kasper for his review of this paper.

REFERENCES

Jacobson, C. M., Kasper, G.M., Mathieu, R.G., McFarland, D.J., and Meservy, R.D. 2011. "AMCIS 2010 Panel Report: External Benchmarks in Information Systems Program Assessment," *Communications of the Association for Information Systems*, (29), Article 20.

Merhout, J., Benamati, J., Rajkumar, T.M., Anderson, P., and Marado, D. 2008. "Implementing Direct and Indirect Assessment in the MIS Curriculum," *Communications of the Association for Information Systems*, (23), Article 24.

Yorke, M. and Zaitseva, E. 2013. "Do cross-sectional student assessment data make a reasonable proxy for longitudinal data?", *Assessment & Evaluation in Higher Education*, (38:8), pp. 957-967, DOI: 10.1080/02602938.2013.769199.