A STUDY ON THE EFFECTIVENESS OF CONTINUOUS ASSESSMENT AS A LEARNING TOOL

W.C.W. Navaratna and M.A.P. De Silva

Department of Mathematics and Computer Science, The Open University of Sri Lanka

INTRODUCTION

Assessing a student is not merely assigning a mark or a grade to reflect the achievement of the learner at the end of studying the course. McMillan (2000) points out that assessment enhances instruction and also influences student motivation and learning. Muirhead (2002) reports that assessment is an important element in the teaching and learning process that challenges instructors to consider evaluation techniques that meet the learning needs of today's adult learners.

Teaching and learning is a continuous ongoing process. It is essential to have ways of assessing the achievements of the students as they proceed. The instructor is left with the challenge of identifying discrete time points at which the assessments should be done and to decide on the kind of assessment strategies that are most appropriate at the selected mile stones.

The Bachelor of Science (B.Sc.) degree programme at the Open University of Sri Lanka (OUSL) is offered as a two-semester program. Each semester is around 15 weeks long. Carefully designed self assessment questions in the course material provide the learner an opportunity to continuously assess themselves as they study the lesson material. Students are more formally assessed through continuous assessment tests (CATs) and end of course final examinations. In the case of mathematics courses offered by the Faculty of Natural Sciences, formal feedback for the learners as well as for the teachers are obtained through two CATs, of which one is an open book test and the other is a no book test. The continuous assessment mark (CAM), computed as a weighted average of the two continuous assessment marks with a weight of 0.6 for the best mark and 0.4 for the other, is used to determine the eligibility to take the end of course final examination. Current practice is to use a CAM of 40 as the eligibility cut off mark that determines the eligibility to take the end of course final examination. The overall final mark for the course is determined as the average mark of the continuous assessment mark and the end of course final examination mark (FEM). A minimum overall mark of 40 is required for a pass grade.

It is worth examining how effectively the students use continuous assessment (CA) as a learning tool. The study described here attempts to address this issue through a careful study of the assessment marks for several mathematics courses offered for the B.Sc. degree programme.

AIM AND OBJECTIVES

The main aim of this study is to examine the effectiveness of CATs as learning tools and to find out how the performance at these tests relate to that of final examinations.

Specific objectives of the study are

- To study the associations of CA marks with the final examinations marks.
- To assess the effectiveness of CA marks in predicting the final examination marks.
- To examine the effectiveness of CATs as learning tools.

METHODOLOGY

The marks pertaining to three Level 3 mathematics courses of the B.Sc. Degree Programme PCU1141, APU1141 and PCU1142 in the academic year 2011/2012 are used for this study.

PCU1141 and APU1141 are identical but PCU1141 is offered by students at Level 04 who are not offering Applied Mathematics as a major discipline as an open elective course. APU1141 is offered by students at Level 03 who are offering Applied Mathematics as a major discipline. PCU1142 is offered by students at Level 04 who are not offering Pure Mathematics or Applied Mathematics as a major discipline.

Scatter plots are used to examine the type of association between CAM and FEM. The strength of the linear association between the CAM and the FEM is measured by the Pearson correlation coefficient (Cox and Hinkly (1974)). The effectiveness of CAM in predicting the FEM is studied through a regression analysis.

To study the effectiveness of CATs as learning tools we need to do a careful analysis of information collected from students on how they feel about the course as well as observations from instructors and also analysis of CAT and FEM marks. In this report, we present the latter, which is, how the performance of students at the final examination had varied depending on the performance at the CATs.

RESULTS AND DISCUSSION

Here we report the results obtained by analyzing the CA marks and final examination marks of PCU1141/APU1141–Basic Statistics and PCU1142–Bio Statistics. Scatter plots of final examination marks against CA marks of all three courses indicated a linear association between the two variables but the points were widely scattered suggesting that the association is not strong. As we already noted only those students who obtain a minimum eligibility mark of 40 are permitted to write the final examination. Therefore, the results we report below are based on the performance of those students. Table 01 summarizes the descriptive statistics and the Pearson correlation coefficient between the CAM and the FEM.

Table 01: Descriptive Statistics

Course Code		Descriptive Statistics						Pearson Correlation Coefficient
		No. of observations	Mean	Median	Standard deviation	Minimum	Maximum	
PCU1141	CAM	53	49.75	49	7.45	40	65	0.213
	FEM	48	54.85	54	10.65	37	86	
APU1141	CAM	104	47.60	45.50	8.36	40	79	0.467
	FEM	78	52.44	52.50	10.86	27	74	
PCU1142	CAM	152	47.15	45	8.046	40	83	0.438
	FEM	116	43.59	40.50	16.16	16	88	

The first column in Table 01 indicates that the number of students who had taken the final examination in each course is less than the number of students who had taken the final examination. This means all students who are eligible to take the end of course final examination have not sat. At the OUSL, once students are eligible for a course, they are permitted to postpone sitting the final examination. For the courses we are examining, the permitted period is three years including the year in which they register.

The second and third columns of Table 01 indicates that for each of CAM and FEM, the sample mean and sample medians are quite close suggesting that the distributions of continuous assessment marks and final examination marks are likely to be symmetrically distributed. Also, the average marks for the two tests in each case are quite similar. However, the standard deviation of the end of course final examination marks for each course is bigger than that for the CA marks. Thus, the students end of course final examinations seem to discriminate the students much more than the CATs. In here we note that CATs are held at

the course at an early stage in the learning process with the objective of giving feedback for learners as well as for teachers. Thus, more rigorous testing which focuses on the cognitive domain is not much deeply tested in the CATs. This is not so for the end of course final examinations. Thus, as one would expect, the final examinations that also focuses on the cognitive domain seem to separate out the students more, thus rendering large standard deviations for final examination marks.

The last column of Table 01 gives the Pearson correlation coefficients that measure the strengths of linear associations between CAM and FEM. In all three cases, the Pearson correlation coefficients are small indicating that the linear associations between the CAM and FEM are not strong. The scatter plots in each case, further justifies this observation. Next we report the findings from fitting a simple linear regression model with CAM as the predictor variable.

Table 02: Regression Analysis output

Course Code	Regression coefficient of CAM (P-value)	Mean squared error	R^2
PCU1141	0.30 (0.145)	110.0	4.5%
APU1141	0.59 (0.00)	93.4	20.8%
PCU1142	0.85 (0.00)	212.8	19.2%

The first column of Table 02 indicates that in the case of PCU1141, the CAM does not significantly contribute to predict the variation in the final examination mark. In APU1141 and PCU1142, the CAM contributes significantly to predict the final examination mark but the last column indicates that CAM alone is not adequate to explain the variation in the final examination marks. We already noted that the scatter plots of data in each case indicate a linear relationship between the two variables. Thus, the failure of CAM alone to adequately describe the variation in FEM can be due to the existence of other factors having a significant effect on the FEM or due to the differences in the effects of CAM as learning tools for different learners. To further study this, we examined the unusual observations highlighted by the regression model fit.

Table 03 presents the number of unusual observations the CAM, FEM and the fitted value for the FEM rounded off to the nearest integer, and the standardized residual (std.resid) for each unusual observation Table 03 highlights information on two types of unusual observations. The observations with large absolute values for standardized residuals correspond to points that do not agree with the fitted linear regression model. It is interesting to note that all of these correspond to students with low CAM marks. In PCU1141 and PCU1142, the large standardized residuals are positive. This means these students at the final examination have performed better than that predicted from the fitted model for their CAM marks. Since both test scores are symmetrically distributed with similar location parameters, this may be due to effectively using CAM as a learning tool. It is worth noting that PCU1141 and PCU1142 are offered by students at Level 04 and they have effectively used CAM as a learning tool. In APU1141, standardized residuals with large absolute values are negative. Thus, those students have performed poorly compared to what is predicted from the simple linear regression model for their CAM. Thus, for them, lower CAM has adversely affected their learning process. APU1141 is offered by students at Level 03. They have not effectively used CAM as a learning tool, but rather this seems to have got an adverse effect.

We also analyzed the CAM of students who were eligible but have not taken the final examination. We find that in PCU1141, APU1141 and PCU1142 9%, 25% and 24% respectively have not sat. Among only those with CAM less than 50, the relevant figures for PCU1141, APU1142 and PCU1142 are 12%, 44% and 31% respectively. Thus, not sitting for the final examination in each course is more prominent among the students with lower CAM, This further emphasizes that the lower CAM has adversely affected the motivation of students to take the final examination further study is needed to confirm this observation.

Table 03: Details on unusual observations

Course	Number of unusual observations	Details on unusual observations			std. residual
		CAM	FEM	Fitted Value	
PCU1141	1	44	86	53	3.20
APU1141	6	40	27	48	-2.16
		47	30	52	-2.27
		49	32	53	-2.18
		68	-	64	-
		68	46	64	-1.94
		79	68	70	-0.28
PCU1142	9	41	73	38	2.40
		46	75	42	2.24
		47	76	43	2.25
		51	80	47	2.29
		64	88	58	2.12
		67	-	60	-
		76	58	68	-0.73
		78	60	70	-0.72
		83	71	74	-0.23

CONCLUSIONS/RECOMMENDATIONS

This study highlighted that in PCU1141 and PCU1142 students with lower CAM have performed better than what is predicted from the model fit. In contrary, in APU1141, students with lower CAM have performed poorly compared to what is predicated based on the CAM. The courses PCU1141 and PCU1142 are offered by students at level 4 whereas APU1141 is offered by students at Level 3. This suggests that students at the start of the degree programme have not effectively used the CA as a learning tool as compared to those at the higher level. Thus, we recommend that it is needed to educate students on how to effectively use CA as a learning tool, especially at the commencement of the program.

The focus of this study is directly linked to improving student performance. This study only focused on the performance of students at the assessment tests. Further study by collecting information from students as well as extending this study to include courses at higher levels are recommended.

REFERENCES

Cox D.R, Hinkly D.V. (1974) Theoretical Statistics Chapman & Hall (Appendix 3) ISBN 0-412-12420.

McMillan J. H. (2000), Fundamental Assessment Principles for Teachers and School Administrators, *Practical Assessment Research and Evaluation*, 7(8).

Murihead B.(2002), Relevant Assessment Strategies for online colleges and Universities, *USDLA Journal*, Vol 16, No. 2.

ACKNOWLEDGMENTS

We are thankful to the Vice Chancellor for giving permission to use the data, referred in this study and Director/IT and SAR/ Exams for making the data available.