**Evaluating Mathematics Competencies of freshmen by diagnostic pre- and post- e-assessments throughout the blended learning mode**

K.H. Chung, Louisa W.M. Wong, C.K. Chan

Department of Applied Mathematics,
The Hong Kong Polytechnic University, Hung Hom, Hong KongEmail: machung@polyu.edu.hk

Abstract

In line with strategic objectives of The Hong Kong Polytechnic University (PolyU), in addressing the major educational development needs of the university due to the adoption of a new four-year university curriculum since September 2012, institutional-wide diagnostic e-assessments and e-learning packages have been developed. This paper describes the extension of this institutional project involving several Departments of the University, conducted by the Department of Applied Mathematics (AMA), mainly on conducting the pre- and post-e-assessments to evaluate freshmen’s performance in blended learning mode and to understand their e-learning patterns. Students taking online pre-diagnostic tests at the beginning of the semester were informed of their performance and provided with e-learning materials accordingly. Towards the end of the first semester, students’ pre- and post-DT scores were compared. Statistical *t*-tests revealed that students had improvement in post-DTs. A multivariate *t*-test and Bonferroni confidence intervals suggested that the improvement was contributed by improvement in differential calculus skill. Data show that students using e-books regularly tend to have noticeable improvement to their post-DTs scores, as compared with those not using the e-books. To further enhance the e-learning materials based upon pedagogical practice, a more in-depth student feedback questionnaire on e-learning materials was adopted in the e-book survey of 2016. Popularity of the e-tools was investigated and general feedback from students was quite positive.

Keywords: Blended learning, pre- and post-diagnostic assessments, e-learning

1. Introduction

The reformed education system of Hong Kong adopted in 2012 has allowed secondary students to take a number of elective subjects and/or extended parts of mathematics for the Hong Kong Diploma of Secondary Education Examination (DSE) in its six-year secondary schools. The extended part of mathematics DSE has two modules: Module 1 (M1) on calculus and statistics and Module 2 (M2) on algebra and calculus. Such changes have admitted university entrants with varying mathematics abilities and provided a huge challenge in mathematics teaching at the university. Under the rapid development of technology, Hoic-Bozic et al. (2009) suggested blended learning support to traditional teaching.

Hung et al. (2015) presented results on institutional-wide diagnostic e-assessments and e-learning packages developed at The Hong Kong Polytechnic University (PolyU) in 2014 to address the major educational development needs since 2012. In particular, the Department of Applied Mathematics (AMA) conducted e-diagnostic assessments and investigated students’ mathematics competencies based on the e-assessments in 2014-15. Ismail et al. (2011) concluded that the self-tests were valuable tools in detecting weaknesses of students, via their experience in pre- and post-diagnostic tests (DTs). This paper extends the work of Hung et al. (2015) in supporting the freshmen of PolyU in learning mathematics, by conducting the pre- and post-diagnostics mathematics tests. Our goal is to allow teachers to identify students’ abilities at the beginning of the semester as well as allowing students to be aware of their strengths and weaknesses in the subject. During the semester, students were provided with the e-learning packages as a self-learning tool. They were encouraged to focus on the topic(s) which they did not perform well in the pre-DTs.

Students who used e-learning packages are called e-book users, while those who had never used the packages are called non-e-book users. Focuses were students’ performance throughout the semester, as well as any gain in blended learning mode in the first semester.

As technology advances, many pedagogical approaches, like dynamic assessments in Wang (2014) and conceptual learning model in Srisawasdi and Panjaburee (2014), were considered in the e-learning platform. To improve the e-learning materials, a more in-depth questionnaire modified from the annual e-book survey was designed and students’ usage of the online e-tools and their comments on the e-books were also collected in the survey of 2016.

2. Methodology

This paper has two parts: evaluating students’ performance via pre- and post- diagnostic e-assessments as well as the e-book survey.

2.1. Pre- and Post- Diagnostic e-Assessments

The pre- and post-diagnostic tests were conducted at the beginning and by the end of the first semester of the academic year 2016-17, in a similar framework as the DT in Hung et al. (2015). The target of PolyU freshmen were from the Faculty of Applied Science and Textile, the Faculty of Engineering and Senior-Year-Admitted students of School of Nursing.

Upon their registration at PolyU, students were invited via emails to take the pre-diagnostic tests. The students were informed that the tests would be used as a tool to identify their strengthens and weaknesses and no special preparation was required for the tests. Two modes of DTs, algebra test (AL test) and algebra & calculus test (AC test), were created in the Learning Management System of Blackboard, LEARN@PolyU. The former test was given to the students without taking Hong Kong DSE M1 or M2, while the latter test was given to the students having taken Hong Kong DSE in Core Mathematics, together with M1 or M2. To encourage the students to take the voluntary diagnostic tests, students were allowed to attempt the tests via LEARN@PolyU at a time and place of their own choice before the end of September 2016. In addition, cake/coffee/ice-cream coupons were provided as incentives to serious test takers.

Participants were screened and those seriously attempted Pre-DT would be informed of details of the post-DTs towards the end of the first semester. The post-tests were held also via LEARN@PolyU from the end of November to the end of December 2016. With similar settings to the pre-tests, the same incentives were provided.

Questions of pre-DT were selected from the question bank for DTs as described in Hung et al. (2015). The AC test consists of questions over four underpinning skills: 10 on algebra, 6 on differential calculus, 8 on integral calculus and 8 on linear algebra. For ease of data analysis, the project team decided to use the same questions for the post-tests as in the pre-tests because the test scores of the two tests could be treated as continuous variables, rather than ordinal ones for analysis.

2.2. Evaluation of e-Packages

Since 2012, three mathematics e-books, namely Foundation Mathematics, Foundation Statistics and Engineering Mathematics, have been made available to students who took mathematics subjects offered by AMA. The e-learning packages enable student-centered independent study outside the classroom and also facilitate group learning, peer assessment and self-assessment. Surveys, targeted at students taking mathematics subjects offered by AMA, were conducted annually to obtain students’ feedback. In spring of 2016, two more questions were added to evaluate the popularity of the e-tools and to obtain the general overall comments about the e-books.

3. Findings of Pre- and Post- Diagnostic e-Assessments

In the pre-DT, 431 participants took the algebra test (AL test), while 329 participants took the algebra and calculus test (AC test). Students’ answering time on each question were recorded in Blackboard. A simple screening rule to identify seriously test takers was used to identify students answering more than 50% of the questions and spending at least half an hour to finish the test. 182 and 146 students were screened as seriously AL and AC test takers respectively. They were invited for the post-DTs and 39 and 35 participants took the AL and AC tests respectively. Outliers were found in the preliminary analysis. They were those taking the test with maximum idle time of at least 30 minutes and were excluded for further analysis. 32 and 30 students were identified as serious AL and AC test takers for analysis.

3.1. Overall Performance

Overall results of both pre- and post-tests scores were summarized in Table 1. Students’ average scores in post-DTs are higher than that of pre-DTs in both tests. Investigation on mathematics competencies improvement was conducted by paired *t*-tests to test whether average post-test score is greater than the average of the pre-test score at 5% significance level. For the AL test group, the sample size is 32, the test statistic is 2.12 and the *p*-value is 0.021, which is less than 0.05. This indicated that the average post-AL test score was statistically significant greater than the average pre-AL test score. For the AC test group, the sample size is 30, the test statistic is 3.21 and the *p*-value is 0.0003. This indicates that the average post-AC test score is statistically significant, greater than the average pre-AC test score. Therefore, it is concluded that the performance of the two test groups has improved after one semester. The mathematical training to the students is effective.

Table 1 Summary of Pre- and Post-AC test and Pre- and Post-AC test Results.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Pre-DT scores | Post-DT scores | DT Scores Difference |
| **Algebra Test Group** (32 students) |
| Mean | 57.7% | 63.8% | 6.1% |
| Standard deviation | 18.9% | 18.3% | 16.4% |
| **Algebra & Calculus Test Group** (30 students) |
| Mean | 57.3% | 62.4% | 5.1% |
| Standard deviation | 18.3% | 15.0% | 8.7% |

3.2. Underpinning Skills of Algebra & Calculus Test Group

The extent of improvement in the four underpinning skills in Algebra (ALG), Differential Calculus (DC), Integral Calculus (IC), Linear Algebra (LA) are shown in Table 2.

Table 2 Summary of Underpinning Skills Performance in Pre- and Post-AC test.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algebra & Calculus Test** (30 students) | Average Pre-DT Scores | Average Post-DT scores | Average Pre- and Post-DT Scores Difference | 95% Bonferroni Confidence Intervals |
| Algebra (ALG) | 88.1% | 88.1% | 0.0% | (-5.8%, 5.8%) |
| Differential Calculus (DC) | 47.8% | 57.2% | 9.4% | (1.0%, 17.8%) |
| Integral Calculus (IC) | 35.8% | 39.2% | 3.3% | (-3.4%, 10.1%) |
| Linear Algebra (LA) | 52.1% | 60.4% | 8.3% | (-2.1%, 18.7%) |
| Overall | 57.3% | 62.4% | 5.1% |  |

A multivariate *t*-test was conducted to test whether there were collectively changes in scores in four skills. The Hotelling’s *T*-square test statistic is 17.8 and the *p*-value is 0.012. Therefore, collectively changes in four skills can be concluded at 5% significance level. To further investigate any improvement in each skill, 95% Bonferroni Confidence Intervals (Johnson and Wichern, 1992) for the mean difference of pre- and post-DT sub-scores in four underpinning skills were presented in Table 2. The confidence interval for DC is in the positive region, while the other intervals cover zero. This indicates the improvement in differential calculus is statistically significant with no statistical difference in the pre- and post-DT scores over algebra, integral calculus and linear algebra. The significant improvement in differential calculus contributes to the overall statistically significant improvement in the AC test.

3.3. Performance Comparison between e-Book Users and Non-e-Book Users

To correlate the student performance with aids of the e-learning packages, students were split into two groups: e-book users and non-e-book users. Their DT scores were also linked with the first semester mathematics subject examination GPA for analysis.

3.3.1 Overall Performance

Figure 1 shows the dot plot of the DT score differences of each student in the two groups, where one dot represents one DT score difference of one participant. For the both tests, more than half of the observations lied in the positive region. This indicates that the e-book users tend to have greater improvement as compared with the non-e-book users.

Figure 1 Distribution of Pre- and Post-DTs Participants’ Scores Differences

Table 3 shows that e-book users spending about 1.5 hours on e-books on average per week and their average DT scores difference are more than 1.5 times greater than that of non-e-book users in both tests groups.

Table 3 Students' Pre- and Post-DTs Performance by e-book users and non-e-book users for AL and AC Tests Groups. \*S.D. refers to standard deviation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | e-book hours per week | Pre-DT Score | Post-DT Score | DT Scores Difference |
| AL Test | e-book users(15 students) | Average | 1.5 | 61.1% | 67.3% | 6.2% |
| S.D. | 2.0 | 15.4% | 20.1% | 17.3% |
| non-e-book users(10 students) | Average | - | 59.7% | 62.6% | 2.9% |
| S.D. | - | 22.4% | 20.1% | 15.7% |
| AC Test | e-book users(10 students) | Average | 1.5 | 62.2% | 69.4% | 7.2% |
| S.D. | 2.1 | 17.9% | 12.0% | 9.8% |
| non-e-book users(12 students) | Average | - | 56.8% | 61.2% | 4.4% |
| S.D. | - | 19.2% | 14.6% | 8.3% |

To test whether the average DT scores difference of e-book users is greater than that of non-e-book users for both DTs, either *t*-test with equal variance or that with unequal variances in two samples was conducted. The equal variance assumption is confirmed by Levene’s Test. When it is the case, pooled standard deviation is computed as an input for the two samples *t*-test. Table 4 shows that results of Levene’s tests and the two samples *t*-tests for both AL and AC tests groups. Both e-book users and non-e-book users groups have the same variances in the pre- and post-DTs score differences and no difference in the improvement between two groups is concluded at 5% significance level for AL and AC tests groups.

Table 4 Testing whether the DT scores difference of e-book users is greater than that of non-e-book users. \*S.D. refers to standard deviation.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Sample size | Levene’s test | Two samples *t*-test |
|  | e-book users | Non-e-book users | Test statistic | *p*-value | Pooled S.D. | Test statistic | *p*-value |
| AL test | 15 | 10 | 0.21 | 0.650 | 16.99% | 0.49 | 0.315 |
| AC test | 10 | 12 | 0.64 | 0.432 | 8.98% | 0.72 | 0.240 |

3.3.2. Underpinning Skills of Algebra & Calculus Test

Average pre- and post-DT scores of the underpinning skills are plotted as a radar diagram in Fig.2. Ten e-book users and twelve non-e-book users exhibit higher post-DT scores in all skills, compared with the pre-DT scores. The extent of improvement in all skills seems to be higher for e-book users than that of non-e-book users. In particular, the extent of improvement in integral calculus (IC) of non-e-book users is smaller than that of the e-book users. This can be confirmed by Bonferroni confidence intervals for mean score difference of e-book and non-e-book users in Table 5. Only the interval for mean integral calculus scores difference does not contain zero and is on the positive side. Therefore, the improvement of integral calculus for e-book users is concluded to a greater extent than that of the non-e-book users.





Figure 2 Average Pre- and Post-DT Scores of Underpinning Skills: Algebra (ALG), Differential Calculus (DC), Integral Calculus (DC) and Linear Algebra (LA) in Algebra & Calculus Test. Here N represents the sample size.

Table 5 Pre- and Post-DT Percentage Scores Difference in Underpinning Skills: Algebra (ALG), Differential Calculus (DC), Integral Calculus (IC) and Linear Algebra (LA) for e-book users and non-e-book users. \* S.D. represents standard deviation.

|  |  |  |
| --- | --- | --- |
| **Algebra and Calculus Test** |  | Post-DT % Score minus Pre-DT % Score |
| ALG | DC | IC | LA |
| e-book users(10 students) | Mean | 3.3% | 8.3% | 5.0% | 13.8% |
| S.D.\* | 14.9% | 18.0% | 15.8% | 17.1% |
| non-e-book users(12 students) | Mean | 2.8% | 5.6% | 1.0% | 10.4% |
| S.D.\* | 10.7% | 17.9% | 15.5% | 20.5% |
| 95% Bonferroni confidence intervals for mean score difference between groups | (-2.2% 3.3%) | (-1.1% ,6.7%) | (0.5% ,7.4%) | (-0.8% ,7.5%) |

**3.3.3. First Semester Mathematics Subject Performance**

Figure 3 shows that all e-book users had GPA of 2.5 or above, while some non-e-book users failed in AL test; and e-book users tended had higher GPA compared with non-e-book users in AC test.

Figure 3 Distribution of Pre- and Post-DTs Participants' First Semester Mathematics Examination GPA: 0 (fail), 1 (D) to 4 (A), with 0.5 increment representing an “+”.

As an examination GPA may refer to different ranges of examination scores for different subjects and GPA has a limited number of ordinal categories of the performance, we have treated the examination GPA as an ordinal variable for analysis. We have adopted Mann-Whitney *U* test (Gay et al., 2009), the nonparametric statistical tool developed based on ranks, on comparing median GPA scores between groups. The null hypothesis is the median examination GPA score of the e-book users is less than or equal to that of the non-e-book users. Table 6 indicates that the *p*-values for both test groups are greater than 0.05. No difference in the results of the first semester mathematics examination between groups is concluded at 5% significance level for both test groups.

Table 6 Mann-Whitney *U* Test on testing whether the median mathematics examination GPA of e-book users is less than that of non-book users in the first semester

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Sample size | Median GPA | Range of GPA | Mann-Whitney *U* test statistic | *p*-value |
| AL Test | e-book users | 13 | 3.5 | 2 | 156 | 0.3443 |
| non-e-book users | 9 | 3.5 | 4.5 |
| AC Test | e-book users | 7 | 4 | 2 | 59 | 0.0874 |
| non-e-book users | 6 | 3 | 1.5 |

3.3.4. e-Book Usage Pattern and First Semester Mathematics Performance

E-book users taken AL test had the median DSE Core Math Score 4 and the range from 4 to 6 (i.e. 5\*), while e-book users taken AC test had the median DSE Core Math score 5 and the range from 4 to 7 (i.e. 5\*\*). This indicates that e-book users are competent learners. Figure 4 shows their first semester Mathematics examination GPA ranging from 2.5 to 4.5 and most of the users spent less than 2 hours on e-books per week. The relationship on exam GPA and e-book usage hours are not obvious.



Figure 4 E-book users' e-book usage pattern and their first semester mathematics examination GPA scores

4. Evaluation of e-Packages

145 students responded to the e-book survey and 73 of them reported that they used e-books. The responses of those e-book users to the two new questions concerning the usage of e-tools, embedded in e-books and general comments over the e-books are plotted in Fig.5 and Fig.6. Figure 5 shows that 46.6%, 60.2%, 54.7% and 42.4% of e-book users reported that they had always, often or sometimes used the interactive e-tools: (a) statistical e-tools, (d) differentiation and integration calculators, (e) graphing e-tool and (g) Gaussian elimination exercise respectively. The most popular e-tool is differentiation and integration calculators. Figure 6 shows that the students’ feedback over the e-books is quite positive. 23 students made suggestions to the e-books. In particular, 9 of them would like to have more exercises in the e-books for self-study.

|  |  |
| --- | --- |
|  |  |
| Figure 5 Students’ usage of (a) statistical e-tools, (b) statistics formulas, (c) statistical distribution tables, (d) differentation and integration calculators, (e) graphing e-tool, (f) calculus formulas and (g) Gaussian elimination exercise | Figure 6 Students’ comments on (a) abundant learning material, (b) appropriate level of materials for their study, (c) adequate exercise and (d) adequate explanation |

5. Conclusion

The pre- and post-DTs results showed that students’ mathematics competencies are statistically significantly improved in the semester. This indicates that the current mathematics teaching by AMA, together with outside classroom support, including e-learning support is effective. For evaluation of effectiveness of blended learning mode, data show that e-book users tend to have better post-test results, as well as higher GPA than non-e-book users. In particular, e-book users improve their integral calculus to a more significant extent than the non-e-book users in the AC test group.

E-books survey revealed that about half of the respondents did not use e-books. Promotion of e-books is required, in particular, the e-tools, which are useful to the independent learners. In addition, adding more exercises to the e-books is the general suggestion made by students. With the small sample size, also containing poor learners, it would be better that the e-assessments are conducted mandatorily under invigilation in order to get the full evaluation of e-support to students.

Acknowledgement

This project was partially funded by a grant from The Hong Kong Polytechnic University (Grant No. LTG12-15/420G). We thank Miss Luo Xin, Tracy for her assistance in the e-package evaluation survey.

References

Gay, L.R., Mills, G.E. and Airasian, P. (2009) *Educational Research: Competencies for Analysis and Applications*. Ninth Edition, Pearson Education.

Hoic-Bozic, N., Mornar, V. and Boticki, I. (2009) “A blended learning approach to course design and implementation”, *IEEE Transactions on Education*, Vol. 52 (No. 1): 19-30.

Hung, K.F., Choy, S.C., Chung, K.H. and Chan, C.K. (2015) “Supporting freshmen through diagnostic e-assessments and e-learning packages”, *Proceedings of the IMA International Conference on Barriers and Enablers to Learning Maths: Enhancing Learning and Teaching for All Learners*. Ed. M.A. Hersh and M. Kotecha, Glasgow, Scotland, U.K., 10-12 June 2015.

Ismail, N.A., Ariff, F.H.M., Asshaari, I., Othman, H. and Nopiah, Z.M. (2011) “Pretesting on the first semester engineering students in UKM: Does it make difference?”, *Procedia Social and Behavioral Sciences*, Vol. 18: 81-86.

Johnson, R.A. and Wichern D.W. (1992) *Applied Multivariate Statistical Analysis*. Third Edition, Prentice-Hall, Upper Saddle River, NJ, U.S.A.

Srisawasdi, N. and Panjaburee, P. (2014) “Technology-Enhanced Learning in Science, Technology, and Mathematics Education: Results on Supporting Student Learning”, *Procedia- Social and Behavioral Sciences*, Vol 116: 946-950.

Wang, T.H. (2014) “Developing an assessment-centered e-Learning system for improving student learning effectiveness”, *Computers & Education*, Vol. 73: 189-203.