



Analysis of quality of test items and students' perception of the online formative tests in Anatomy

Nilesh Kumar Mitra, Umesh Bindal, Shajan Koshy

ABSTRACT

Objectives: This study was planned to assess the quality of test items of online tests in Anatomy using quiz statistics indices of Moodle software and to compare the perception of the students of two consecutive cohorts towards the online formative tests in Anatomy. **Methods:** The study was conducted among two consecutive cohorts (n=45) of first year medical students in a medical school in Malaysia. Five online MCQ tests on different Anatomy topics were given during the 16 weeks time of semester 1. Using quiz report statistics, mean grade, facility index (FI, difficulty level) and discrimination index (DI) of the test items were computed. At the end of the semester, the students were given questionnaire to assess their perception. A descriptive cross sectional design was adopted, frequencies were calculated followed by crosstabulations and analyzed with Chi-square distribution test. **Results:** Except one test, mean FI of all the tests remained in the acceptable range of 30 to 80%. Mean grade in the consecutive tests changed proportional to the increase or decrease of the ease of the test items. Mean DI was found to be in acceptable range (0.3 or more) only in two tests in first cohort and three tests in second cohort. The study found 62% and 61% of students in first and second cohort agreed that online tests helped them to understand position of preparation in Anatomy. No significant difference was observed between the perceptions of students of two consecutive cohorts towards the online test helping their preparation in Anatomy (Chisquare test, $P=0.092$). **Conclusion:** Automated quiz statistics of Moodle software can be used to assess the difficulty and discrimination levels of the test items in online formative tests. The study also reflected a positive perception of the majority of year 1 medical students towards the online formative assessment in Anatomy.

KEY WORDS: Quiz statistics, Formative tests, Moodle, Facility Index, Discrimination Index, Students' perception

Department of Anatomy, School of Medicine, Taylor's University, Jalan Taylor's, 47500, Subang Jaya, Selangor, Malaysia

Address for correspondence:
Nilesh Kumar Mitra, Department of Anatomy, School of Medicine, Taylor's University, Malaysia,
NileshKumar.Mitra@taylors.edu.my

Received: September 29, 2015

Accepted: October 22, 2015

Published: November 03, 2015

INTRODUCTION

The pedagogy of teaching-learning in the medical education has been changed with the advent of e-learning. Assessments are formative if they motivate students to learn and direct their learning process [1]. Tutors need to know the deficit in the students' knowledge in order to place special emphasis on providing feedback on areas in the subject matter [2]. Introduction of web-based and multiple-choice-question (MCQ) based formative assessment embedded in online Learning-Management-System (LMS) provide the students with ready-made feedback depending on their response. Moodle is the most popular open-sourced LMS in the world. According to a recent market research by Capterra, a software guidance company, Moodle tops list of the "20 most popular LMS software solutions" [3]. Out of twenty six Malaysian universities, twelve used Moodle as the LMS [4,5]. Moodle's quiz module has automated statistical methods to measure the reliability of the tests [6]. Psychometric analysis like Difficulty Index or Facility Index (FI) and Discrimination Index (DI) can help to find out whether the questions are of an appropriate level of difficulty, and are suitable enough to discriminate between good and bad performers [7].

Buchanan (2000) suggested that the strategies in feedback like 'repeat the test' and 'timely feedback' greatly benefit learning effectiveness among college students [8]. The

student's perception of the learning environment determines how he or she learns [9]. The absence of the teacher during the interaction of the students with the web-based online formative assessment necessitates understanding of the students' perception towards the assessment process, the environment and the outcome. The objectives of this study were to apply the quiz statistics indices of online formative tests in Anatomy given to two consecutive cohorts of semester 1 students of MBBS program to assess the quality of the process and compare the perception of the students of two consecutive cohorts towards the online formative tests in Anatomy.

METHODS

The study was conducted among the first year students of the MBBS program in the School of Medicine of Taylor's University, Malaysia. Two consecutive cohorts of students (n=45) admitted to MBBS program during March 2013 and August 2013 intake, named as batch 2013-1 and batch 2013-2 respectively, were selected for this study. The study was given ethical clearance from the institutional ethics committee. Response to the questionnaire collected from the students did not include name or any other identification of the students. Anatomy curriculum in Semester 1 Foundation Block module has significant basic learning

objectives like General Anatomy, Histology, Embryology and Musculoskeletal system. The topics are taught via didactic lectures, practical and students' seminar. The continuous paper-based summative assessments are held on 7th week, 14th week and at the end of the module. Five online Multiple-Choice-Question (MCQ) tests having 15 items were given during the 16 weeks time interspersed across the Semester 1 module. The students were briefed that the marks would not be used for summative assessment. The topics covered in the test were chosen based on the topics taught in the previous 3 weeks (i.e., General Anatomy/Histology, General Embryology, Upper Limb, Lower Limb, and Neck). The MCQ items used in the formative tests were not used in the continuous summative assessment.

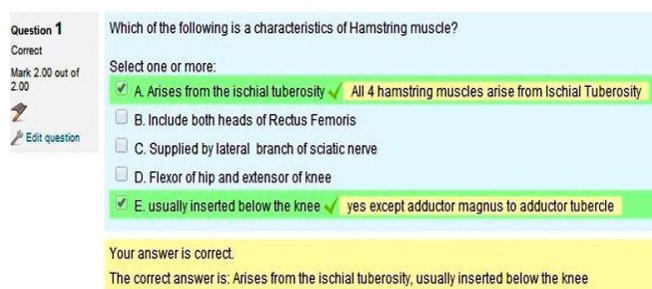


Figure 1. Sample of feedback received by the student after attempt of the online quiz item

Students did not get the correct answer during the first attempt. They were allowed to attempt the quiz after 48 hours again and get the automated feedback according to their response (the option ticked by them) (Figure 1). For quiz report statistics, performance at 1st attempt was computed. To meet the first objective, the quiz report statistics of all the five tests given in batch 2013-1 and batch 2013-2 were generated to find out mean grade of class in each test and FI and DI indices of each item of all tests. A Facility index (FI) measures how easy or difficult is the question item. The system computes it as $FI = X \text{ average} / X \text{ max}$ (ratio of mean credit obtained by all users and maximum credit achieved in the item). FI value below 30% is categorised as difficult and above 80% is categorised as easy question. Discrimination Index (DI) indicates the proficiency of the item to discriminate between high-achievers and low-achievers. The system computes it as $DI = (X \text{ top} - X \text{ bottom}) / n$ (Difference between mean credit obtained by upper group and lower group divided by number of students taking the test). Test items with minus DI are very poor discriminator and should be removed. Items with DI of 0.3 or more were accepted to be able to discriminate between high-achieving students and low-achieving students [10]. The data of the assessment quality indices for batch 2013-1 and batch 2013-2 were compared using descriptive statistics of SPSS version 17.

To meet the second objective, a descriptive cross sectional design was adopted and a questionnaire was developed for

data collection. At the end of the semester 1 module, the students were given the questionnaire. The first half of the questionnaire focused on obtaining feedback about their competency in use of computer, operating system of the Moodle software and the interface of the examination. The rest of the questionnaire contained questions on whether the quiz item covered the outcome of the module adequately and whether automated feedback after response helped them in the learning process. For each question, students were given a Likert scale, to select one of the five options (strongly disagree, disagree, somewhat agree, agree, strongly agree). Some of the relevant statements in the questionnaire were taken from a study conducted in United Kingdom [11]. The data were analyzed using SPSS version 17, initially using frequencies followed by cross tabulations to generate descriptive statistics. Statistical analysis was conducted using the Chi-square distribution test.

RESULTS

Analysis of quiz statistics

The mean grade of the class at first attempt, as shown in quiz statistics improved from the online formative test 1 to formative test 5 (Figure 2). Both cohorts of students showed initial average poor performance in the first online formative test on General Anatomy/Histology (mean class grade of 35% in batch 2013-1 and 28% in batch 2013-2). Consecutively the performance improved in subsequent tests. The difficulty level as expressed by mean FI (Figure 3) of the consecutive online formative tests remained in the acceptable range of 30 to 80% except in the third test of batch 2013-1. The increase of the index above 80% signified that the test items were easy and FI 30% or below should reflect that items were getting difficult. Figure 4 showed the comparison between mean FI of the test and the mean grade of the class. The left sided scatter chart of performance of batch 2013-1 followed identical skew of mean grade of class with increasing ease of test items in 2nd and 3rd online tests and decreasing ease of test items in 4th and 5th online tests. Batch 2013-2 performance (right sided scatter chart) did not show similar change in performance. Although the mean grade in batch 2013-2 showed proportional increase with MCQ items being easier with increasing facility index in test 1 and 2, such relationship was not followed in test 3, 4 and 5. Out of total test items used, 57% showed acceptable DI (0.3 or more) in batch 2013-1 and 68% showed acceptable DI in batch 2013-2. The mean DI of the online tests was found to be in acceptable range only in two tests in batch 2013-1 and three tests in batch 2013-2 (Figure 5). In both batches of students, with increasing ease of test items above facility index of 70%, there was observable decrease of discrimination index below 0.2. Test items which were comparatively easier could not discriminate well between high-achieving and low-achieving students

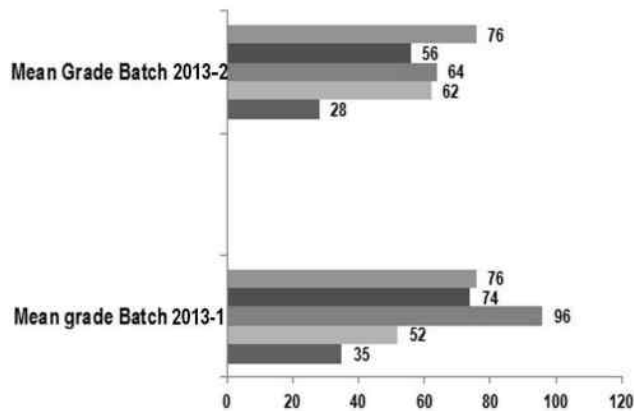


Figure 2. Bar chart showing the mean grade of class in both batches in the formative tests

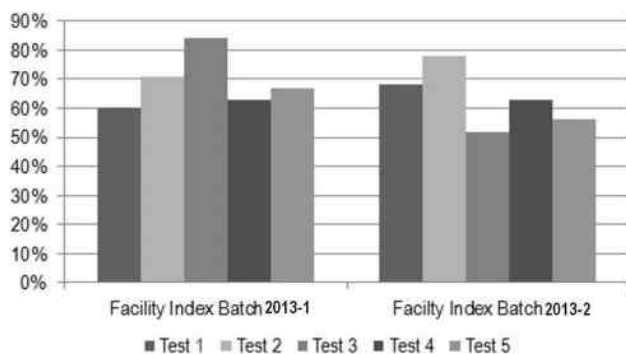


Figure 3. Bar chart showing mean difficulty level of the online formative tests as Facility Index (in percentage) in the quiz statistics of the moodle software

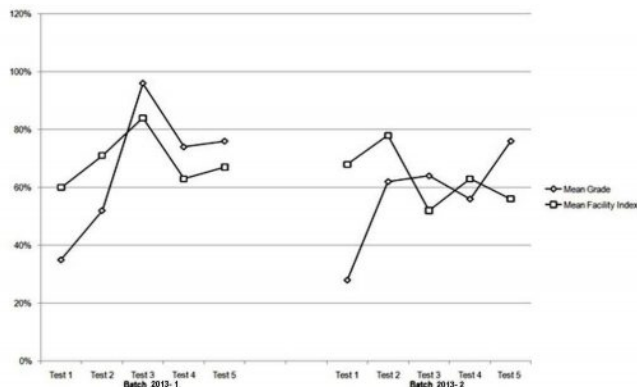


Figure 4. Scatter chart of mean grade of class in percentage compared with that of mean Facility Index (difficulty level) of each formative test

Perception of the students about the online formative tests

The response rate to the questionnaires distributed at the end of the semester was 87%. Majority of the students in both cohorts of students believed that their competencies with computers was just enough (46% in batch 2013-1 and 43% in batch 2013-2) (Figure 6). A good computer competency

was claimed by only 38% students in batch 2013-1 and 34% in students in batch 2013-2. Experience of previous online assessment was found in 97% of the students in batch 2013-1 and 86% in batch 2013-2. While 10% of students in batch 2013-1 did not agree at all that the interface was clear, only 2% of batch 2013-2 students were found with similar opinion. When asked whether the online test helped the students to understand their position of preparation in Anatomy, 36% strongly agreed and 26% agreed in batch 2013-1. With similar question, out of batch 2013-2 students, 32% strongly agreed and 29% agreed that they could understand their position of preparation in Anatomy with the online formative tests (Figure 7). No significant difference was observed in the distribution of agreement, disagreement and neutral perception between students of batch 1/13 and batch 2/13 about whether the online test helps them to understand their position of preparation in Anatomy (Chi-square test, $P=0.092$). The number of students expressing disagreement with the above-mentioned question was 10% in batch 2013-1 only. The strategy of multiple attempts and option-specific feedback were planned so that the online test would help the students to learn from the mistake while undertaking the test. In batch 2013-1, 54% agreed that the feedback received after attempt of each item helped them in their learning process. With the improvement in the feedback system in next cohort, 70% of batch 2013-2 students agreed that the feedback received helped them in their learning process.

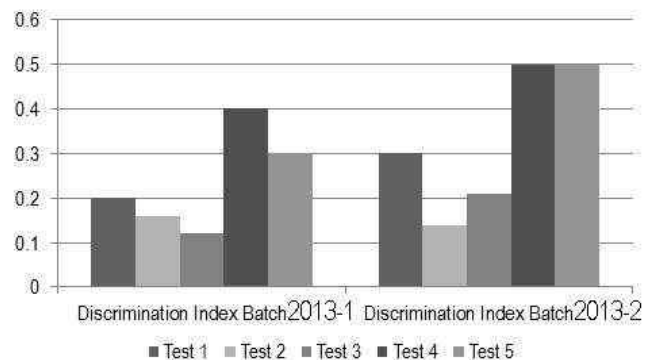


Figure 5. Bar chart showing mean Discrimination Index of the online formative tests from the quiz statistics of the moodle software

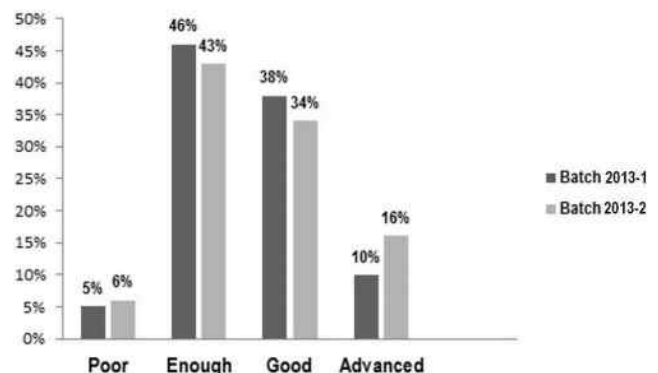


Figure 6. Bar Chart showing students' own perception of competency in computing skill

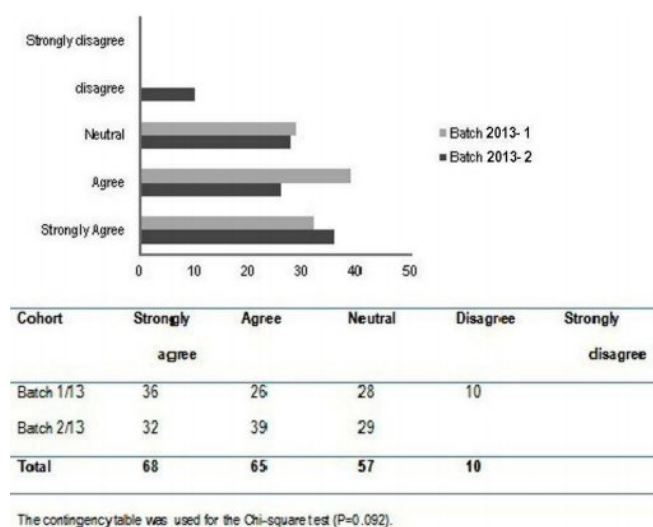


Figure 7. Distribution of perception among students (in percentage) about whether online formative tests help in understanding their preparation in Anatomy

DISCUSSION

A computer-based online formative assessment can be self-administered and the student can assess himself or herself without the score being known to other students [12]. Computer-based assessment with a feedback process is beneficial for the learning process as it provides immediate feedback [13]. The students' performance in the form of mean grade changed proportional to the increase or decrease of the ease of the test items (corresponding to increasing or decreasing facility index) except last few tests in batch 2013-2. The last two tests for batch 2013-2 had very good discrimination index (0.5) which might have altered the proportional relation of mean grade of class with the facility index. For both cohorts of students, discrimination ability of the test item was found to be decreased once the difficulty level of the item decreased with increasing facility index above 70%. Two previous studies have also found similar relationship of reduction of discrimination ability of test items with increasing difficulty index [10, 14].

Only one-third of the students in both batches admitted that they had good computer competency. Wallace and Clariana (2000) found that the learners who were less familiar with content and less familiar with computer did not do well in online assessment [15]. When the students were asked about their perceptions of learning, students with surface approach, deep approach and strategic approach behaved differently [16]. It is expected that students with surface approach would feel it as a burden to go and prepare for online formative test every week in addition to attending usual lectures, practical and seminars. Two-third of the students agreed that online test helped them to understand their position of preparation in Anatomy. The statistical analysis found that there was no significant

difference between the number of students with agreement, disagreement and neutral perception between both batches of the students. Providing feedback facilitating learning and tag it to online test environment in Moodle software was a learning experience for the Anatomy lecturers. Increase in the percentage of students, in agreement to the statement that the feedback received after attempt of each item helped them in their learning process, from 54% in first batch to 70% in second batch of students, signified some improvement in the feedback process.

LIMITATIONS

The study was conducted in consecutive 2 cohorts of medical students from Malaysian population after the online formatives tests with feedback was implemented in the Anatomy curriculum. The feedback from the lecturers about the online tests was also not analysed. The limited data might deter the generalisation of the findings to some extent.

CONCLUSION

With the increasing number of students in the medical schools, online formative assessment can be an option to facilitate the learning process in Anatomy curriculum, in which huge contents are required to be covered within limited number of semesters. The results of this study have shown that use of Moodle software with automated quiz statistics for these online tests can help the teachers to assess the difficulty and discrimination levels of the test items with a view to improve the quality of test items. The study has also reflected a positive perception of the majority of student population towards the online formative assessment.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

1. Leung C, Mohan B. Teacher formative assessment and talk in classroom contexts: assessment as discourse and assessment of discourse. *Lang Test* 2004;21(3):335-59.
2. Black P, William D. Assessment and classroom learning. *Assess Educ Princ Pol Pract* 1998;5(1):7-74.
3. Moodle tops list of the 20 most popular LMS software solutions. In: Moodle News Available via <http://www.moodlenews.com/2013/moodle-tops-list-of-the-20-most-popular-lms-software-solutions/> [Accessed 28 September 2015].
4. Hamat A, Embi MA, Sulaiman AH. Learning management systems in Malaysian higher education Institutions. In: *E-Learning in Malaysian higher education institutions: status, trends and challenges*. Department of Higher Education, Malaysia, pp 30, 2011.
5. Juhary J. The learning management system at the defence university: awareness and application. *Int Educ Stud* 2013;6(8)e-pub.
6. Ferrão M. E-assessment within the Bologna paradigm: evidence from Portugal. *Assess Eval High Educ* 2010;35(7):819-30.
7. Blanco AM, Ginovart GM. On how moodle quizzes can contribute to the formative e-assessment of first-year engineering students in mathematics courses. *Univer Knowl Soc J (RUSC)* 2012;9:354-70. Available via <http://rusc.uoc.edu/index.php/rusc/article/view/v9n1-blanco-ginovart> [Accessed 28 September 2015]

8. Buchanan T. The efficacy of a world-wide web mediated formative assessment. *J Comp Assist Learn* 2000;16(3):193–200.
9. Entwistle N J. Approaches to learning and perceptions of the learning environment. Introduction to the special issue, *High Educ* 1991;22(3):201–4.
10. Mitra NK, Nagaraja HS, Ponnudurai G, Judson JP. The levels of difficulty and discrimination indices in type A multiple choice questions of pre-clinical semester 1 multidisciplinary summative tests. *Int e-Journal of Sc Med Edu* 2009;3(1):2-7.
11. Lilley M, Barker T. Students' perceived usefulness of formative feedback for a computer-adaptive test. *The Electron J Elearn* 2007;5(1):31–8.
12. Pitt SJ, Gunn A. The value of computer based formative assessment in undergraduate biological science teaching. *Biosci Educ Electron J* 2004;3e-pub
13. Brown S, Knight P. Assessing learners in higher education. RoutledgeFalmer, pp 102, 1994.
14. Sim SM, Rasiah RI. Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper. *Ann Acad Med Singapore* 2006;35(2):67-71.
15. Wallace PE, Clariana RB. Achievement predictors for a computer-applications module delivered via the world-wide web. *J Inform Syst Educ* 2000;11(1/2):13-18.
16. Struyven K, Dochy F, Janssens S. Students' perceptions about evaluation and assessment in higher education: a review. *Assess Eval High Educ* 2005;30(4):331–47.

© SAGEYA. This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, noncommercial use, distribution and reproduction in any medium, provided the work is properly cited.
Source of Support: Nil, Conflict of Interest: None declared