Harnessing Learning Analytics to Improve Online Quiz Equity

Jess Wei Chin Tan, Singapore University of Social Sciences, Singapore
Chong Hui Tan, Singapore University of Social Sciences, Singapore
Hian Chye Koh, Singapore University of Social Sciences, Singapore

The Asian Conference on Education 2019
Official Conference Proceedings

Abstract
The use of online learning management systems (LMS) such as Blackboard, Canvas and Moodle is becoming a norm in higher education. These systems have been used widely for student assessments such as online quizzes. The Singapore University of Social Sciences (SUSS) established in April 2005 (then known as SIM University), also uses pre-class online quizzes to encourage students to self-study. The questions used for the quizzes are usually randomly drawn from a 50-question bank. The level of difficulty for the questions varies and this raises the question of quiz equity. This paper investigates how the current method of question allocation affected the equity of the quiz. It also proposes a solution to mitigate the issue of quiz inequity. With an integration of learning analytics and problem solving, we hope to provide a different approach to implementing online quizzes that will be more equitable.

Keywords: Online quizzes, random assignment, quiz equity, learning analytics.

1. Introduction

The use of online learning management systems (LMS) such as Blackboard, Canvas and Moodle is becoming a norm in higher education. In general, these systems provide tools to facilitate active learning such as discussion forums and student assessments such as online quizzes. Studies (Johnson and Kiviniemi (2009), Narloch, Garbin, and Turnage (2006), Salas-Morera, Arauzo-Azofra and García-Hernández (2012)) have shown that students who do pre-class online quizzes that encourage preparatory reading perform better in examinations. Dobson (2008) explains that these quizzes motivate students to be better prepared for class through the reading of course materials such as textbooks, study guides and notes before class, which ultimately leads to more effective learning of the course materials.

The Singapore University of Social Sciences (SUSS), established in April 2005 (then known as SIM University), is a university that caters primarily to working adults. Its mission of lifelong education equips learners to serve society as well as contribute to the workforce. In line with self-directed lifelong learning, it uses pre-class online quizzes to encourage students to self-study. SUSS uses Canvas as the learning management system to house and implement these quizzes. The pre-class quizzes constitute a small percentage of the students' overall grade (less than or equal to 6%). In particular, the students need to obtain at least 12 out of 20 questions correct for the first pre-class quiz (also termed as pre-course quiz), which is available to them at least one week before class starts. Otherwise, they will not be allowed to attend classes. This further motivates students to engage the course materials before the classes start. The faculty will usually prepare a bank of questions from which the questions for the pre-class quizzes will be randomly drawn for each student. The level of difficulty for the questions varies. This raises the issue of quiz equity despite the random allocation of quiz questions.

In this context, this study investigates how the current method of question allocation has affected the equity of the quiz. It also proposes a solution to mitigate quiz inequity. With an integration of learning analytics and problem solving, we hope to provide a different approach to implementing online quizzes that will be more equitable.

2. Literature Review

While there does not seem to be any prior investigation of the control and effect of the selection of quiz questions, issues of learning analytics and pedagogy design have been well studied. This helps to provide the background for our study, for which the principal thrust concerns the application of learning analytics to improve the design and implementation of pedagogy and the related artefacts for online learning.

Learning analytics is an emerging discipline that lies in the crossroad of data analysis and pedagogy (Siemens (2013)). The discipline focuses on a better understanding of how data produced and collected from the learning process can be fruitfully analysed in educational settings. Siemens (2013) has traced the historical developments of the field through the enumeration of conferences, journals, summer institutes and research labs to conclude that learning analytics has strongly emerged as an important research area. Thus, consistent with this development, the setting up of the Society for Learning Analytics Research (SoLAR) has connected international researchers on the impact of analytics on teaching, learning, training and development since 2011. SoLAR also has its own publication, the Journal of Learning Analytics. In a landmark study, Ferguson (2012) has charted the developing areas of learning analytics research and challenges.

Martin and Ndoye (2016) apply learning analytics to assess student learning in online courses. Through a case study, they discover and affirm that learning analytics data on student activities in the learning management system helps guide the implementation of courses, design of assessments and the provision of support to weaker students.

Tempelaar, Heck, Cuypers, van der Kooij, and van de Vrie (2013) classify information sources for learning analytics into two main types – intentional and learner-activity metadata. Intentional data arises from formative assessments retrieved from learning management systems, while learner-activity metadata records facets of learner interaction on the systems. In an integrated infrastructure, Tempelaar, Heck, Cuypers, van der Kooij, and van de Vrie (2013) combine both types of data to provide feedback to students and teachers on personal learning dispositions, attitudes and values.

At Purdue University, Arnold and Pistilli (2012) rely on learning analytics in the form of grades, demographic characteristics, past academic history and students' effort to provide feedback to students. The goal is to help students improve their performance by keeping them informed about their academic position in a course in real time. The project, called Course Signals, is reported to have been deployed to over 23,000 students across 100 courses and 140 instructors.

It is evident from these studies that learning analytics hold promise for positive intervention in the teaching and learning process at a multitude of interfaces and in a variety of situations. Our present work involves the real-time collection of learning analytics data for improving the online quiz design through more equitable selection of question sets. With reference to the literature on online quizzes, this falls into a spectrum of concerns that can be classified into research on the pedagogical effects and effectiveness of online quizzes (Dobson (2008), Angus and Watson (2009)) and quiz design considerations (Gray and Jackson (2003), Kajitori, Aoki and Ito (2014)). The former delves into the existence of positive impacts on learning or examination performance through statistical tests performed on data sets obtained from online quizzes. On the other hand, the latter covers issues of operational effectiveness through the use of technology such as round-the-clock access (which promotes self-learning in a safe environment), automatic generation of hints and the inclusion of a retry functionality which improve the learning experience.

3. Method

In this study, we analysed the pre-course quiz of an introductory business analytics course delivered in a blended mode to 300 students in July 2018. Students are required to obtain at least 12 out of 20 questions correct for this quiz before they are allowed to attend classes. The 20 questions are randomly drawn from a 50-question bank. The pre-course quiz is available to the students at least one week prior to the first class and students are allowed to make unlimited attempts during the period that the quiz is made available to them. Only the best score will be recorded and form part of the students' overall grade (≈2%). The students' performance of the pre-course quiz is very important as it determines whether the students can proceed to attend classes. This also raises the question of quiz equity that might affect the students' performance of the pre-course quiz.

Canvas provides quiz statistics at the student- and question-level, with all attempts recorded. The difficulty index, tabulated as part of the quiz statistics, indicates how hard it is to answer a particular question correctly. The index is computed as the proportion of correct answers over total number of answers and hence is reverse-scored (i.e., the higher the index, the more the correct answers, which in turn means the easier the particular question). A scatterplot of the mean (x-axis) and standard deviation (y-axis) of the difficulty index of the set of questions taken by each student and the students' quiz performance for the first (and completed) attempt (of all 20 questions) is as shown in Fig. 1. The first and completed attempt was used in the analysis as a means to eliminate attempts that might constitute students trying to "game the system".

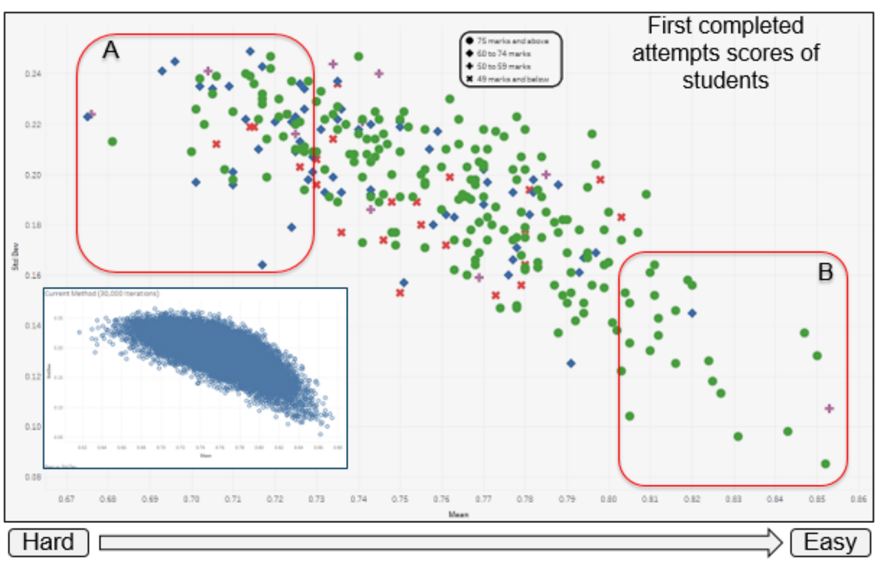


Figure 1. Random allocation of questions and quiz score.

It can be observed that the current allocation approach of 20 questions randomly drawn from a 50-question bank has resulted in some students getting proportionately more questions that are either easy or difficult. In particular, students in Box A have been allocated proportionately more of the harder questions and have performed relatively poorer. The converse is true for students in Box B. The insert is a plot based on 30,000 Monte Carlo simulation iterations of quiz question allocation. It shows that this inequity is inherent in the current method of quiz question allocation. To mitigate this, a new quiz-question random-allocation method that considers the difficulty level of each question is proposed. Two methods are presented in Fig. 2A and 2B below.

The Optimal Method bins (i.e., groups) questions based on "clean breaks" in their difficulty index (e.g., Easy, Fairly Easy, Fairly Hard, and Hard) as shown in Fig. 2A.

Figure 2A. Binning of questions based on the optimal method.

On the other hand, the Feasible Method bins questions based on different quartile-ranges of their difficulty index (e.g., 0%-25%, 26%-50%, 51%-75%, and 76%-100%). This is shown in Fig. 2B.

Figure 2B. Binning of questions based on the feasible method.

For each of the proposed methods:

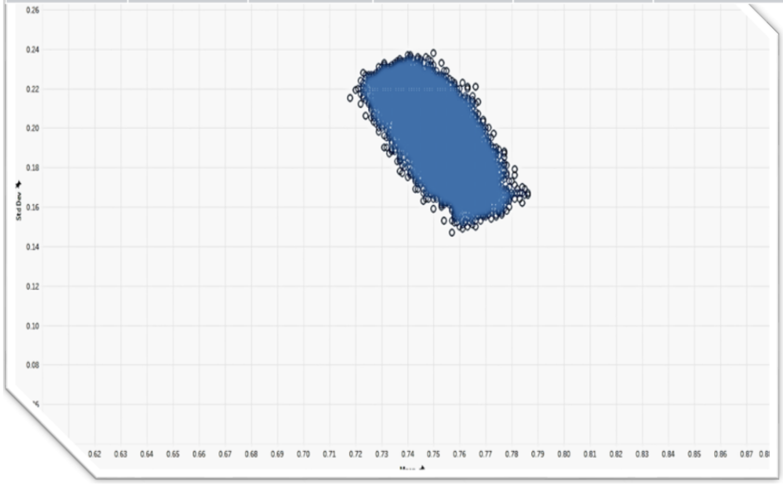
- All the questions are first grouped into bins by difficulty index, and then questions are randomly selected from each bin in appropriate proportions to make up the required number of pre-course questions, which is 20 in this study.
- For validation, 30,000 Monte Carlo simulation iterations are generated and compared with the outcomes of the current method of allocation.

4. Findings

Simulating 30,000 random draws of 20-questions sets using the optimal and feasible methods, while adhering to the stipulations as discussed, produce distributions in the question-set characteristics as illustrated in Fig. 3.

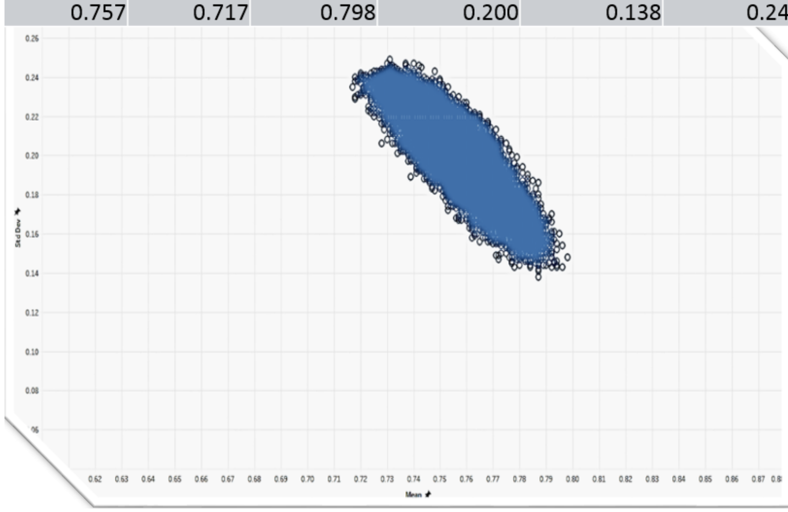
Optimal Method:

Ave(Mean)	Min(Mean)	Max(Mean)	Ave(Std Dev)	Min(Std Dev)	Max(Std Dev)
0.752	0.718	0.786	0.196	0.147	0.238



Feasible Method:

Ave(Mean)	Min(Mean)	Max(Mean)	Ave(Std Dev)	Min(Std Dev)	Max(Std Dev)
0.757	0.717	0.798	0.200	0.138	0.249



Current Method:

Ave(Mean)	Min(Mean)	Max(Mean)	Ave(Std Dev)	Min(Std Dev)	Max(Std Dev)
0.758	0.617	0.874	0.194	0.055	0.265

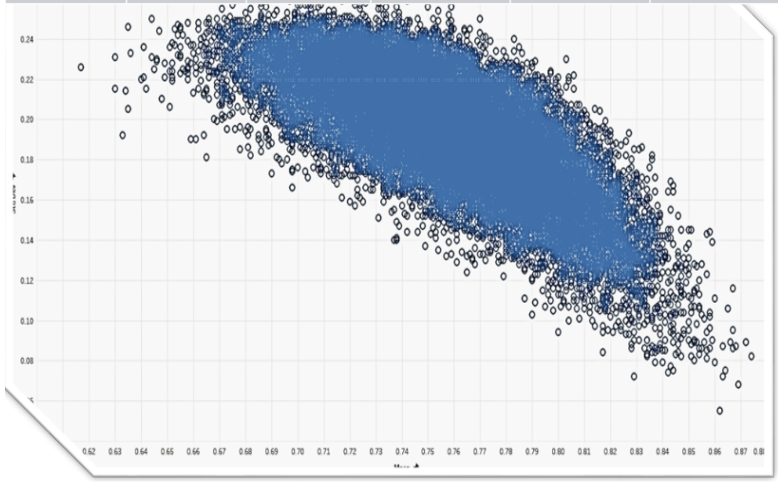


Figure 3. Distribution of questions for the optimal and feasible method (versus the current method (based on 30,000 iterations)).

Fig. 3. shows very clearly that the ranges for the mean and standard deviation of the difficulty index for both the optimal method and the feasible method, as compared to the current method, have reduced significantly. This implies that the question-set generated using either optimal/feasible method have smaller variations in difficulty across the quizzes vis-à-vis the current method.

The ranges (i.e. max – min) of the mean of the difficulty index in both the optimal method and the feasible method have reduced significantly to 0.068 (0.786-0.718) and 0.081 (0.798-0.717), respectively, while it has been 0.257 (0.874-0.617) under the current method. This signifies that the spread of average difficulty of a question-set has been reduced under either of the proposed methods thereby increasing quiz equity.

The standard deviation of a quiz indicates the variability of question difficulty within it. A high standard deviation (which is preferred) indicates that a quiz has a good variety of questions of differing difficulty (instead of consistently easy questions or consistently difficult questions). However, a large range of these standard deviations across the quizzes is not preferred as this implies the lack of consistency across quizzes in terms of the distribution of easy and difficult quiz questions within individual quizzes, which in turn, adversely affects quiz equity. Hence, within a quiz a high standard deviation is preferred, whereas across quizzes a small range of the standard deviations is preferred. The ranges of the standard deviation of the difficulty index for both the optimal method and the feasible method have also reduced significantly to 0.091 (0.238-0.147) and 0.111 (0.249-0.138) as compared to the current method of 0.210 (0.265-0.055). This indicates that there is greater consistency in the allocation of quiz questions of varying difficulty level across quizzes under either of the proposed methods vis-à-vis the current method. Quiz equity is consequently enhanced.

Finally, we conducted a Chi-Square Goodness-of-Fit test to establish that the distributions of quiz based on mean and standard deviation have significantly departed (i.e., improved) from the distribution obtained with the current method (Fig. 5). To do this, the observations of both optimal and feasible methods are partitioned into 4 buckets (known as the expected values) and compared with the observations of the current method (based on and known as the observed values). The partition of the 4 buckets is based on the mean and standard deviation of the quiz difficulty index. In particular, Bucket 1 represents question sets that comprise mainly difficult questions whereas Bucket 4 contains question sets that comprise mainly easy questions. Bucket 2 and 3 are more desirable in term of quiz equity as both buckets contain a good mix of difficult and easy questions. Collectively, the outcomes in the quizzes of both buckets are within an acceptable range of quiz difficulty.

The test shows that the null hypothesis can be rejected, i.e., there is a significant improvement in quiz equity (i.e., (concentration of question sets in bucket 2 and 3) using the optimal/feasible method vis-à-vis the current method. Collectively, this signifies that both the optimal and feasible method are able to significantly affect quiz attributes to enhance quiz equity.

Optimal Method:

Bin

	Observed N	Expected N	Residual
1.00	96	42.0	54.0
2.00	39	92.0	-53.0
3.00	35	135.0	-100.0
4.00	130	31.0	99.0
Total	300		

Test Statistics

	Bin
Chi-Square	490.197[a]
df	3
Asymp. Sig.	.000

Feasible Method:

Bin

	Observed N	Expected N	Residual
1.00	96	10.0	86.0
2.00	74	188.0	-114.0
3.00	23	60.0	-37.0
4.00	107	42.0	65.0
Total	300		

Test Statistics

	Bin
Chi-Square	932.140[a]
df	3
Asymp. Sig.	.000

Figure 5. Chi-Square Goodness-of-Fit Test: Current method versus Optimal/Feasible method.

5. Conclusion

This study is an integration of learning analytics and problem solving that examines random allocation of questions for online quizzes. Based on the findings, it can be concluded that proportionate random assignment by difficulty bin can significantly enhance quiz equity. The optimal method is also deemed to be more favourable in terms of the findings; however, this binning of the questions requires human judgement. To ease deployment, it is recommended that the feasible method be used as it can be fully-automated (vis-à-vis the optimal method). The difficulty index for existing quiz questions can be extracted directly from the learning management system, while that for new quiz questions can be decided by quiz writers.

This recommendation is generalisable to all courses, although this study is based only on an introductory business analytics course. This applies to the extent that there is a substantial range of question difficulty level and there is no consideration of this difficulty level in the random allocation of questions.

It is hoped that this study provides an efficient and effective approach to implementing online quizzes that will be more equitable.

References

Angus, S. D., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology,* 40(2), 255-272.

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 267-270). ACM.

Dobson, J. L. (2008). The use of formative online quizzes to enhance class preparation and scores on summative exams. *Advances in Physiology Education,* 32(4), 297-302.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.

Gray, J., & Jackson, C. (2004). The development of an online quiz for drug calculations. *Nursing times*, 100(4), 40-41.

Narloch, R., Garbin, C. P., & Turnage, K. D. (2006). Benefits of prelecture quizzes. *Teaching of Psychology,* 33(2), 109-112.

Johnson, B. C., & Kiviniemi, M. T. (2009). The effect of online chapter quizzes on exam performance in an undergraduate social psychology course. *Teaching of Psychology,* 36(1), 33-37.

Kajitori, K., Aoki, K., & Ito, S. (2014). Developing a Compact and Practical Online Quiz System. *International Journal of Modern Education and Computer Science,* 6(9), 1.

Martin, F., & Ndoye, A. (2016). Using learning analytics to assess student learning in online courses. *Journal of University Teaching & Learning Practice,* 13(3), 7.

Salas-Morera, L., Arauzo-Azofra, A., & García-Hernández, L. (2012). Analysis of online quizzes as a teaching and assessment tool. *Journal of Technology and Science Education,* 2(1), 39-45.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist,* 57(10), 1380-1400.

Tempelaar, D. T., Heck, A., Cuypers, H., van der Kooij, H., & van de Vrie, E. (2013, April). Formative assessment and learning analytics. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 205-209). ACM.

Contact email:
Jess Wei Chin Tan: jesstanwc@suss.edu.sg
Chong Hui Tan: chtan@suss.edu.sg
Hian Chye Koh: hckoh@suss.edu.sg