

Proposal of a Conceptual Framework for an efficient and effective Test Management System

TickMeng Lim, Open University Malaysia
limtm@oum.edu.my

Abstract: A test management system is programmed to store and manage test items as well as to automate the generation of tests for the purpose of assessments. An item bank is a repository of test items with associated data, and it is the core component of a test management system. From administrative point of view, the primary focus in the design of a test management system is system efficiency, both in terms of effort and cost. Educators however are more concerned with the validity and reliability of tests in measuring students' performance. Thus, the design and development of a quality item banking system should take both views into consideration. Other issues in the item banking design include test objectivity, appropriateness of item distribution, fairness and practicality. In this paper, the author identifies the essential elements required of an effective and efficient item banking system. Some of the important factors to be considered in designing the system include: the application of Bloom's Taxonomy in item categorization, the role of test specification table in item selection, the capability of the system in filtering poor quality items for review and enhancement. Based on the analysis, a conceptual framework for the design of the item banking system is proposed. The system aims towards minimizing the manual process of the test paper preparation whilst ensuring the quality of the test paper generated. The author provides a detailed description of the framework as well the process involved in test items generation and test items enhancement.

INTRODUCTION

Formal learning has been traditionally perceived as a form of planned structured learning carried out in group settings at an institution situated in a particular locality. Such learning is constrained by the factors of time, space and locality. With the increasing calls for democratization of education and life-long learning, the notion of having learners learning under one roof with instructional delivery from the instructor to the learners is no more relevant and practical in meeting the needs of large and diverse populations of learners who are so widely distributed, not only within the country, but also include other parts of the world. This is even more so with the advent of open and distance learning (ODL). With the setting up of ODL institutions, learning is no more a factor of locality, time and space. In theory, there is no limit to the population size of

students in any ODL institution. Nevertheless, the population size of any ODL institution is still somewhat limited by the capability of the institutions in managing the learning and assessments activities of the students.

As institutions of learning grow in size, the task of conducting formal assessments for students becomes laborious and cost intensive. Handling large number of students and numerous courses make administrative processes of assessment and evaluation a nightmare for ODL institutions (OKONKWO, 2008). Taking the case of Open University Malaysia (OUM) as an example, in every semester, the examination department of the university needs to manage and conduct assessment examinations for more than 20000 students at more than 60 learning centers all over Malaysia. The number of examination papers in any one semester exceeds 500. Since OUM students are so vastly distributed, the administration of the examination process at the end of every semester is not only tedious, laborious and time consuming but also costly. The process includes but not limited to the following stages:

- Identifying and engaging various qualified subject matter experts (SMEs) for the preparation of examination questions and marking schemes
- Reviewing examination questions and marking schemes submitted by qualified reviewers
- Printing and delivery of the examination papers to various examination centers
- Administering examinations at all centers according to stipulated schedule
- Collecting and sending back the answer scripts to the examination department

Considering the many stages and the many people involved in handling the various tasks, it is rather difficult to find a full-proof strategy to avoid any form of leakage in the process. Another issue raised by the academics is: since in each semester, new sets of examinations papers are set, not necessarily by the same subject matter experts, how would we be able to ascertain consistency in the quality of the examination papers prepared?

The use of technology such as item banking and test management system have been regarded by many researchers as the possible solution to solving the problems related to administering tests and examinations mentioned above. In this paper, the author identifies and analyses the essential elements required of an effective and efficient test management system. Based on the analysis, a conceptual framework for the design of the test management system is proposed. The system aims towards minimizing the manual process of the examination papers preparation but at the same time ensuring the quality of the examination paper generated. The author provides a detailed description of the framework as well the process involved in test items generation and test items quality management.

ITEM, BANKING AND TEST MANAGEMENT SYSTEM

Item banking is not a new innovation. It has long been advocated as a possible tool for the development of effective and efficient tests and examinations (Choppin, 1976). In the early years of its development, item banking system appears to be more of a database system to support the mass storage and easy selection and retrieval of items used as examination questions (Estecs, 1985). Little emphasis has been placed on automation of generation and test quality control process. To distinguish the difference between the traditional item banking system with the proposed system and for the purpose of discussion in this paper, the author defines an item banking as just the examination questions storage system which represents a component of the proposed test management system. A test management system here is defined as a comprehensive computer-based system that enables examination items entry, selection, quality management, examination generation and scores analysis,

TEST RELIABILITY AND TEST VALIDITY

In OUM, the main purpose for conducting examinations is to obtain a measure of students achievements against the course objectives. SMEs who are involved in preparing the examinations questions need to ensure the suitability of the items in terms of validity, reliability and level of difficulty.

Test validity refers to whether a test items measure what they intend to measure. It is critical to ensure test validity because if a test lacks validity, then the test scores may not be valid or functional in segregating students according to the actual performance. Test validity, particularly the content validity, are typically determined by the SMEs themselves. In any particular examination, it is almost impossible to have test items covering all aspects of learning outcomes. One way to ensure content validity is to have a good distribution of questions that are representative of the whole module (Jandaghi & Shaterian, 2008).

Test reliability refers to the consistency of the test. Test reliability is a factor of test items reliability. A test item is reliable if students who have better grasp of the subject content scores better in the item as compared to students who have less understanding of the content. Test reliability is often regarded as a precursor to test validity. In other words, if test items included in an examination paper are not reliable, then it is impossible to conclude that the examination validly measures students' level of understanding of the area to be tested.

It is indeed necessary to ensure test validity and reliability as they are both critical factors. Nevertheless it is indeed cost ineffective to undergo stringent process in preparing valid and

reliable test for use in just one specific examination. Besides test validity and test reliability, another aspect of concern is the suitability of a test of examination in terms distribution of items in terms of levels of difficulty. Also, relying on the SMEs to determine the difficulty level of a question based on his own perception may not be an objective way. Fortunately, with current technology, it is now possible for us to carry out various forms of thorough and accurate item analyses to monitor item reliability and test validity.

TEST SPECIFICATION TABLE

A test specification table is often regarded as a blueprint that guides in the preparation of examination questions. It is in the form of a table that displays the distribution of the examination questions for a given subject according to topics to be tested and the cognitive level of questions. In planning assessments based on the traditional mode of delivery, the distribution of questions in the test specification table also reflects what has been taught in the instructional sequence. However, in ODL mode of learning like OUM, self-learning using the print module constitutes 80% of notional learning time. As such, any blueprint used to guide assessment should be heavily based on the content of the learning module. In other words, instead of matching test items to what is taught in the classroom to ensure content validity, the focus is now placed on ensuring that the test items are representative of the content in the learning module.

Besides the module content, another dimension that we have to consider in building a table of specification is the distribution of items according to the levels of cognition. Bloom's Taxonomy which categorizes questions according to six levels of cognition is quite popularly used for the purpose. The six levels are: knowledge, comprehension, application, analysis, synthesis and evaluation respectively.

The Faculty of Education and Languages (FEL) in OUM adopts a specific structure for its examination (as shown in Figure 1). A typical examination paper has three parts. Part A consists of 5 questions with short answers. The questions are aimed measuring learners' learning outcomes at knowledge and comprehension levels. Part B consists of 3 structured essay questions. The questions are aimed at measuring learners' learning outcomes at application and analysis levels. Part C contains 2 essay questions aimed at assessing learners' learning outcomes at synthesis and evaluation levels. The questions are prepared based on the content in the print modules, and a typical print module is divided into 10 topics.

Topic	Cognitive Levels of Questions						No of Questions
	Know-ledge	Compre-hension	Appli-cation	Ana-lysis	Syn-thesis	Eva-luation	
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
							10

Figure 1: Examination Structure in FEL

Currently, SMEs who are instructed to prepare a test specification table prior to preparing the examination questions. The test specification table is of the same structure as in Figure 1. The purpose of building the test specification table is to ensure that the test questions are well distributed in terms of topics and levels of cognition. Nevertheless, it is difficult to ensure that all SMEs understand and truly apply the principles of good distribution when filling-in the test specification table. With technology, it is now possible to automate a test specification table generate the distribution of question types based on pre-set criteria.

PROPOSED FRAMEWORK OF A TEST MANAGEMENT SYSTEM

A test management system should be designed with the main objective of reducing the laborious manual process of examination or test items preparation and administration but at the same time, ensure the quality of examination papers prepared. To achieve the objective, it is of the opinion that the design of an effective and efficient test management system should have the following functionalities:

- i) An item bank which serves as repository to store examination questions according to topics as well as levels of cognition. The basic structure of the proposed item bank for OUM is shown in Figure 2.
- ii) A computer-generated test specification table which determines the distribution of the types of examination questions to be retrieved based on the preset criteria. An example of the generated table of specification is shown in Figure 3.
- iii) An engine to retrieve examination questions and marking schemes (a set of metadata) from the item bank based on the information obtained from the test specification table
- iv) An engine to capture students' examination scores (another set of metadata) from OUM's Online Marks Entry System (OMES) for each question
- v) An engine to perform continuous analysis of the questions based on the captured scores to provide information (such as difficulty index, discrimination index and reliability index) about the quality of the question
- vi) An engine that identify and report on questions that need to be reviewed (based on the readings of the indices)

TOPIC	Know- ledge (C1)	Compre- hension (C2)	Appli- cation (C3)	Ana- lysis (C4)	Syn- thesis (C5)	Eva- luation (C6)
T1	T1C1	T1C2	T1C3	T1C4	T1C5	T1C6
T2	T2C1	T2C2	T2C3	T2C4	T2C5	T2C6
T3	T3C1	T3C2	T3C3	T3C4	T3C5	T3C6
T4	T4C1	T4C2	T4C3	T4C4	T4C5	T4C6
T5	T5C1	T5C2	T5C3	T5C4	T5C5	T5C6
T6	T6C1	T6C2	T6C3	T6C4	T6C5	T6C6
T7	T7C1	T7C2	T7C3	T7C4	T7C5	T7C6
T8	T8C1	T8C2	T8C3	T8C4	T8C5	T8C6
T9	T9C1	T9C2	T9C3	T9C4	T9C5	T9C6
T10	T10C1	T10C2	T10C3	T10C4	T10C5	T10C6

Figure 2: Basic structure of proposed Item Bank

Test Specification Table							
Module Code: HBMT3103		Semester: May			Year: 2012		
Taxonomy Level	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Total
Topic							
1	1						1
2			1				1
3		1		1			2
4							0
5			1				1
6		1				1	2
7		1					1
8							0
9	1						1
10					1		1
PART A (5 Questions))	2	3					5
PART B (3 Questions))			2	1			3
PART C (2 Questions))					1	1	2

Figure 3: A system-generated Test Specification Table

Basically the proposed test management system has two major components. The first component deals with quality management and the second component deals with examination papers generation.

The quality management component has three sub-systems. These sub-systems are, the item banking system, examination items analysis system. The item banking system stores and categorizes examinations questions prepared by SMEs and reviewed by the reviewers. It also stratifies questions based on their respective difficulty indices and discrimination indices generated by the examination items analysis system. The examination item analysis system analyses the items based on students scores collected from the Online Marks Entry System currently used the OUM. It generates report for problem questions that need to be reviewed by the reviewers before depositing into the data bank system again.

The examination papers generation system also has three main sub-systems. The first sub-system automates the generation of test specification tables. Based on the information provided by the test specification table generated, another sub-system retrieves the examination questions that meet the criteria in a random manner. The questions are then formatted into printable form.

A graphical representation of the proposed test management system framework is shown in Figure 4.

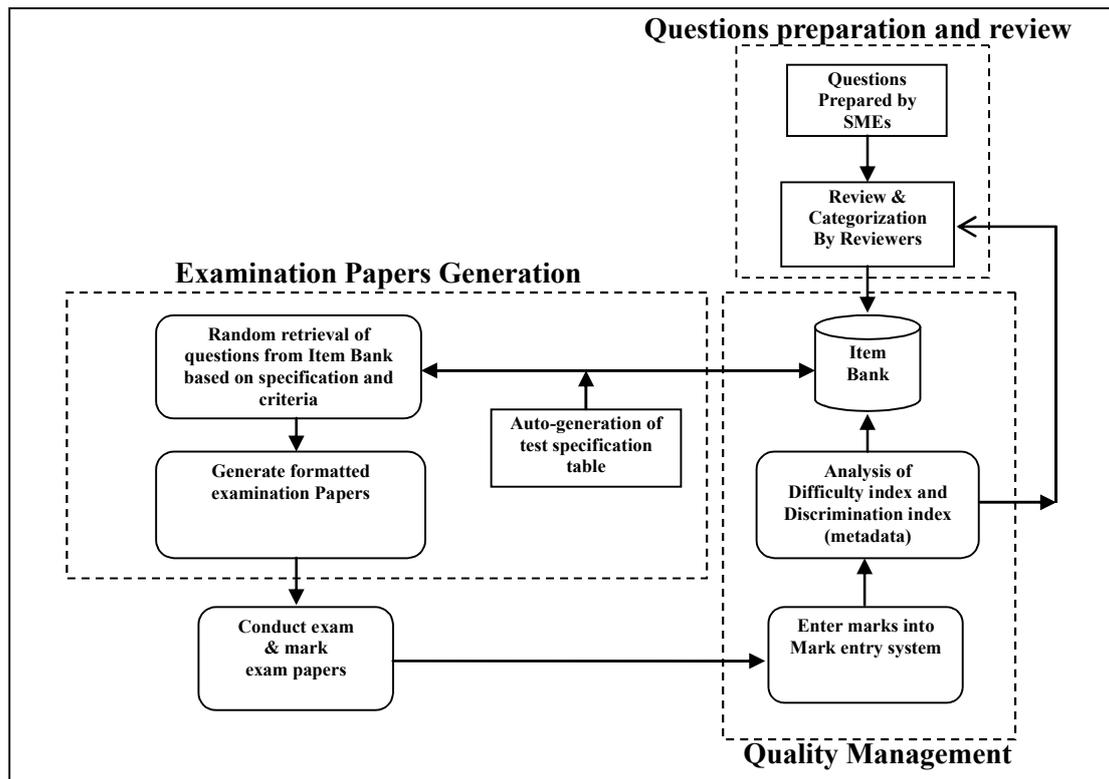


Figure 4: Proposed Framework for OUM's Test Management System

IMPLEMENTATION ISSUES

Ideally, the design of a test management system should aim towards full automation. But there are a number of problems to overcome before full automation can be realized. Firstly, it is necessary to have large enough database of items for every category of items. Secondly, the items need to undergo sufficient review and after-test analysis process to ensure the quality.

Another possible obstacle to full-automation is the format of the examination papers conducted. Currently, item banking systems used by institutions of learning appear to be focused on the "objective kind" of items with fixed answers, (Dolia, 2012), used for the purpose of conducting online evaluation. However, OUM currently adopt the 5-3-2 format for its undergraduate course examinations which comprises 5 questions with short answers, 3 essay questions structured into parts, and 2 long essay questions. Since the examinations items are mainly open-ended questions, it may be technically difficult at this stage to automate the marking of the examination papers in OUM.

CONCLUSION

ODL is probably one of the best possible approaches to make education accessible to everyone. Technology however is a determining factor for ensuring success in implementing ODL. It is in fact the enabler that facilitates the transformation of learning from traditional mode to ODL mode. If technology is appropriately and optimally used, there is a possibility for ODL to replace traditional education as the most preferred choice of learning. One of the areas of administration where the optimal use of technology should be thoroughly explored is the area of assessment. An efficient implementation of test management system will probably benefit not only the institution, but also its learners. It is therefore worthwhile for ODL institutions which have not optimized technology for managing assessment to consider such form of technological empowerment. The proposed conceptual framework represents a small step towards the dream of a fully automated teaching, learning and assessment environment.

References:

Choppin, B. (1976, June 1976). *Developments in item banking*. Paper presented at the First European Contact Workshop, Windsor, UK.

Dolia, P. M. (2012). Direction to prepare item bank for the purpose of online evaluation. *International Journal of Modern Education and Computer Science (MECS)*, 4(5), 18-24.

Estes, G. D. (Ed.). (1985). *Examples of item banks to support local test development: Two case studies with reactions*. Washington, DC: National Institute of Education.

Okonkwo, C. A. (2010). Sustainable assessment and evaluation strategies for open and distance learning. *Turkish Online Journal of Distance education - TOJDE*, 11(4), 121-129.