THE ETEST BUILDER: A RENAISSANCE IN EDUCATIONAL ASSESSMENT PRACTICE

Gereo Alviola Patrimonio
Silliman University (PHILIPPINES)

Abstract

eAssessment, the integration of gauging learning paradigms with computer technology, is an innovative eLearning tool. As a strategy for a computerized educational measurement, it has set the trend for a systemic reform in an educational ecosystem that has been plagued for the last 10 years with problems on inadequate digital literacy. The Silliman University College of Nursing (SUCN) in central Philippines is no exception to this reality. Challenged to uphold the Commission on Higher Education’s (CHED’s) seal of excellence in nursing education, and with the University’s mandate to integrate eLearning as a strategy in the delivery of service, SUCN implores to harness the benefits of eAssessment despite the Nursing faculty’s lack of expertise in the digital realm. This automated test builder, equipped with features that include automated test banking, valid recycling, and item analysis are innovating educational measurement for nursing schools. It was developed compliant to the tenets in Systems Development Life Cycle (SDLC) employing Rapid Application Development (RAD) as methodology with Evolutionary Prototyping technique. A Users’ Training session was also conducted to the Nursing faculty members that exposed them to the automated system. The software quality involved a survey on Nursing faculty members using adapted questionnaires from the International Organization for Standardization-International Electrotechnical Commission (ISO/IEC) 9126-1 framework. The survey yielded a 100 percent valid response rate and that the test bank module is flawlessly-robust and highly-secure. The test questionnaire-generation module’s evaluation revealed that it is absolutely-timely and highly accurate. The statistical report-generation module was considered by SUCN as an excellent decision-making tool. The software product quality evaluation of the eTest Builder, as defined by ISO/IEC 9126-1, showed that the computed aggregate means for functionality, reliability, usability, efficiency, and security is highly-functional, approvingly-reliable, predominantly usable, exceptionally-efficient, and highly secure. Overall, the respondents perceived the eTest Builder as a highly-valuable business solution to the nagging assessment issues and concerns.

Keywords: e-learning Tool, Electronic Assessment, Assessment Instructional Material in Context, Educational Technology, Electronic Test Maker, Software Product Quality.

1 INTRODUCTION

Instruction has three major components: a) objectives which outline the learning to be gained; b) activities which are the actions and events to happen in the course of delivering the instruction; and c) assessment which is the process of measuring what has been learned, that consequently verifies if the objectives have been activated, along with the effectiveness of the activities [1].

Test construction, on the other hand, is governed by rules and sound techniques to ensure validity and reliability. This discipline is not imposed solely to develop the skills required, but for academicians to exercise an appreciation and critical thinking of their strengths and weaknesses [2].

The proliferation and widespread use of computer technology in the early 80s have influenced irreversible changes in the domain of educational measurement. These gave rise to the conceptualization of computerized educational measurements, also known as computer-based assessment (CBA) and electronic assessment (eAssessment). A computerized educational measurement is the generic label for any assessment activity that makes use of the computer technology [3].

The United Nations Educational, Scientific, and Cultural Organization (UNESCO), through its Institute for Information Technologies in Education (IITE), has acknowledged the significant value of technology-enhanced assessment (TEA) in the academe. In its May 2011 Policy Brief, UNESCO-IITE boldly pushed for a computer-supported or fully online testing, as a sophisticated but flexible tool to assess digital literacy. The realization was anchored on a 2006 report from the Programme for International Student Assessment (PISA) that conducted a computer-based assessment method in its
Science Survey. In January 2012, UNESCO-IITE formally endorsed through a Policy Brief the use of TEA in Education. The brief reiterated the need for an innovative tool through a computerized educational measurement “in the context of a systemic reform of the educational system” [4].

Silliman’s strategic plan clearly spells out the importance of eLearning curricular integration [5]. While imbued with the academic seal of excellence in Nursing Education, the Silliman University College of Nursing (SUCN) takes on the formidable responsibility of maintaining its superior national performance rating in the Philippine National Licensure Examination (PNLE).

Cognizant of the tenets of test construction, major examination items that are developed based on actual professional Board questions tend to simulate a real Board examination, reducing the anxiety and increasing the confidence of exam-takers. Cassady and Johnson found this to be true in their study in which high levels of cognitive test anxieties were associated with significantly lower scholastic aptitude test scores [6].

Results of another study conducted in 2012 proved a high anticipation for the development and deployment of an eTest Generator as eLearning tool among faculty members of SUCN hence, its development is recommended [7]. SUCN is apprehensive to Board-like type of questions because of their complexities including, but not limited to: a) technical competencies on the maintenance of item bank that includes the calibration of items to qualify in the test bank employing item analysis; b) the integrity and validity in recycling test questions; c) the tedious design of clinical cases or scenarios as basis for the test items; d) the promotion on the possibility of team-teaching; and e) the establishment of security measures to safeguard the confidentiality of the exam, and to ensure its integrity.

The study seeks to give a business solution through the development of an eTest Builder as an eLearning tool for SUCN, and for the nursing education in general. Embarking to harness the benefits of ICT for the nursing education, this study is geared towards the development of a web-based test questionnaire builder for SUCN. The system will be called eTest Builder, a term coined by the researcher to refer to an electronic test builder, that will be used by the SUCN faculty to comply with the mandate of the university’s strategic plan on eLearning.

2 SYSTEM SCOPE

The eTest Builder is a web-based system geared towards promoting team-teaching, underscoring collaborative effort in the design and construction of clinical cases or scenarios, and corresponding test items by subject area.

Operationally, an eTest Builder is referred to as an online information system consisting of a database of questions, choices, and corresponding correct answers for specific subject matters. The database of questions is calibrated using item analysis. Among other notable features, incorporating intelligence to the system, the eTest Builder will allow test questions to be recycled with another set of corresponding correct answers. At the outset, it shall promote team-teaching and the collaborative effort of generating sets of test questionnaire for a given testing period that is equipped with features for proper monitoring, command responsibility, team effort, and security [7].

Specifically, the system development focused on three general business processes:

2.1 Test Banking Module

This module maintained the pool of items to be banked with the following key points:

2.1.1 Robust test bank

This allowed the creation of clinical cases or scenarios that can be classified according to subject, topic, or theme, equipped with the log mechanism for multi- or single-authorship, date and time stamps creation, alteration, publication, and frequency of use among others.

2.1.2 Valid recycling

Used a mechanism that qualified recycling and allowed for the reuse of a question but generating a new answer key in the choices.
2.1.3 Compliance to assessment regulations
Implemented a weight scheme in each item derived from the Item Analysis procedure and the categories of the Bloom’s Taxonomy as basis for the simple-to-complex sequence in the generated test questionnaire.

2.1.4 Compliance to the Rules of Test Construction
Implemented a Help scheme to show as guide the Rules of Test Construction when populating the item bank.

2.1.5 Collaboration platform
This built-in feature has allowed users of the system to collaborate for the streamlining of clinical cases or scenarios and test items.

2.1.6 Integration of administrative functions
The implementation of the login mechanism among faculty members has rationalised command-responsibility issues and allowed them to exercise leadership in streamlining the item bank maintenance and test questionnaire generation.

2.2 Test Questionnaire Generation Module
This module implemented business procedures that empowered each team through the automated identification of test items in a questionnaire, promoted team-teaching, and compelled them to generate effective test questionnaires that were guided by two key points:

2.2.1 Two-fold agreement mechanism
This mechanism allowed the team-teach leaders and members to agree on the coverage of test through the use of the Table of Specification (TOS) and the items to be included in the test questionnaire to be generated.

2.2.2 Automated test item selection
This mechanism generated the test questionnaire based on the agreed TOS by the team-teach groups and the adherence to the progressive levels of difficulties in the test questionnaire.

2.3 Enterprise Reporting
This module generally aided management in the decision-making endeavour to streamline assessment procedures of the College implementing two key points:

2.3.1 Interactive querying
This feature, as inspired by the online analytical processing (OLAP), has allowed users to analyse multidimensional data interactively under varied circumstances.

2.3.2 Data visualization
This mechanism has facilitated in the clear communication of information through the use of information graphics derived from statistical treatment of data in the test bank.

3 RELATED LITERATURE

3.1 Theoretical Framework
Fig. 1 depicts a multi-dimensional framework of theories serving as the backbone for this study.

3.1.1 Assessment theory
Assessment, as Laurie Brady and Kerry Kennedy discuss it, comes in different forms – short answer test, essay, multiple choice test, quizzes, and oral presentation. They may differ in functions but all
these forms are aimed at gauging a variety of learning. This learning encapsulates the knowledge and skills the students acquired [8].

Assessment also serves as an evaluative tool for the effectiveness of the instruction strategy. It provides timely information to allow the remapping and reconstructing of strategies that will help the student master his subject better. It can also be used to rank academic institutions, and serve as basis for comparison of programs, with reference to the quality teaching-learning procedure that are practiced, reflective of the curriculum used. More so, the disaggregated data culled from an assessment can serve as basis to identify the content areas that will point to group of struggling students [9]. Curriculum planners and implementers in streamlining the educational system's design therefore can use assessment results as basis.

The scholastic performance of a student is drawn from the results of an administered assessment. It reflects the sum total of his understanding, preparation, rehearsal, retention, degree of attention equated to the subject, and his ability to analyse, create, and synthesize from the experienced instruction. This holistic approach is anchored in the tenets of a 1965 framework on Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain by the Committee of College and University Examiners headed by Benjamin S. Bloom [10].

The contexts of validity, reliability, and norms are crucial in the role of measurement in instruction that academicians must be knowledgeable about and cognizant in developing assessment tools that are deliberate and purposive. The concept of validity is the extent of use that fulfills the evaluative task. And for the purposes of this study, validity refers to the quality of the questions being cogent, logically-sound and factually-correct. Likewise, the concept of reliability is the extent of purpose for which the evaluative tool is used. Hence, reliability in this study refers to the consistency of the questions being definitive, well-founded, and established truths. Similarly, the concept of norms involves the established and standardized statistics of the assessment tool describing the evaluative performance [2].

In a public forum on “Assessment of Student Learning,” it was learned that item analysis can be completed using either qualitative or quantitative approach. The qualitative approach utilizes the judgement of experts to meet the requirements of the tool used.

The quantitative approach, on the other hand, employs statistical computations, the results of which are used as basis for judging the desirability or the undesirability of a particular item. These results fill in the variables that would measure an item’s quality or utility in terms of its difficulty index; discriminating power index; and measure of attractiveness of each of the distracters (in case of a multiple choice format). This project only made use of the difficulty index and the discriminating power index because of the introduction of multiple keys in the stem to qualify valid recycling strategies.

The difficulty index ($D_i$) refers to the proportion of the number of students in the upper and lower groups who answered an item correctly hence, also referred to as the item difficulty or item easiness.
The discrimination index \( (D_i) \) on the other hand, serves as an indicator of each item's efficacy in drawing the distinction between students who know it like the back of their palm from those who have no knowledge on the matter. It is also known as the item effect that figuratively, separates the sheep from the goats. The item’s discrimination index is a point-biserial correlation coefficient with the possible values ranging from -1.00 to +1.00. A positive and close-to-one index indicates a better discrimination power. The correlation would somehow indicate that students who got the item correctly have a propensity to get a high overall score. An item with a negative to zero (-1 to 0) index would indicate undesirability.

The values of the item analysis, difficulty and the discrimination indices, are only computed once with the values stored in the test bank. These values form part of the selection of items for the generation of test questionnaires, and comply with the simple-to-complex sequence of items as required in the assessment [11].

3.1.2 eAssessment

Remarkably, more studies have emerged on effective instructions and assessment in the context of promoting higher order thinking competencies of students. Some findings pointed to the curriculum’s inadequacy of preparing the graduates for the professional world; that there is a relative disparity between current assessment practices, and the aspirations of the constituencies [12].

Another study showed that in the practice and adherence to the concepts in assessment, what were initially identified as ideal were not being practiced in reality. Armed with a “Next Generation Learning” paradigm, the authors acknowledged and sought the aid in information and communications technology (ICT) to achieve its goals. They prescribed, “how new thinking and technologies can help” [13].

Literatures of eAssessment mostly point to the Item Banking Model and Item Response Theory (IRT) as its forerunners. Item banking is grounded on psychometric basis with mathematical equations as its foundation [14]. Over the years, the model of IRT has been translated in a number of assessment procedures administered by legitimate bodies around the globe. The Automated Item Selection (AIS) procedure developed by Swanson and Stockings has been used by the Educational Testing Service (ETS) for the administration and scoring of Test of English as a Foreign Language or widely known as TOEFL [15].

The International Association for Computerized Adaptive Testing or IACAT, an organization for scientific, educational, literary, and charitable purposes dedicated for the scientific advancement of adaptive testing applied to psychological and educational measurement [16], is also a recipient of the enormous scientific research and development of IRT as engine for the innovative CAT. CAT is the “redesign of psychological and educational measuring instruments for delivery by interactive computers.” Used to test a person’s ability, CAT operates through the use of pre-calibrated item bank, manipulated using IRT to effectively and efficiently measures a person's traits [17].

Another widely known entity in the field of nursing that heavily uses CAT is the National Council of State Boards of Nursing (NCSBN) of the USA. The two licensure examinations maintained by NCSBN are the National Council Licensure Examination for Registered Nurses (NCLEX-RN); and the National Council Licensure Examination for Practical Nurses (NCLEX-PN). Both licensure examinations are administered and scored using CAT. The National Council of State Boards in the US heavily depends on the effectiveness of ICT in carrying out security measures, thereby reducing security risks in the testing procedure [18].

4 THE SYSTEM DEVELOPMENT

4.1 The Functional User Requirements

What the system should accomplish in terms of the services and the constraints on its operations are reflective of what the clients’ needs. Functional user requirements are declaration of services available to the users that covers how the system reacts on particular input and behaves on specific situation [19]. Using the defined scope of the eTest Builder as guide, the series of rapid application development (RAD) workshops successfully captured the functional user requirements of the proposed system. The eTest Builder shall be able to:

1. enroll all faculty members of the SUCN as main users of the system;
allow a registered user to have control over his own private information;
3 have access levels to establish and to facilitate the delegation of tasks and accountability to maintain the integrity of the stored data and system's performance;
4 integrate the SUCN curriculum for the nursing subjects as major reference to generate test questionnaires;
5 incorporate the topics under each nursing subjects as anchor to determine the coverage of the test questionnaire;
6 provide appropriate user views to support collaboration platform;
7 apply and integrate the concepts of the Table of Specifications (TOS) and the Bloom’s Revised Taxonomy on the Classification of Educational Objectives for Learning Outcomes in the course of determining the coverage of the test questionnaire;
8 enroll semestral team-teach groups that handle the nursing subjects offered in a particular semester;
9 build the different components of a cell – clinical cases or scenarios, items with keys and options that populate the test bank;
10 provide built-in mechanisms that allow comments and actions for approval for each contributed cell, test questionnaire topic outline, and generated test questionnaire;
11 perform item analysis procedure for a particular generated and administered test questionnaire;
12 provide appropriate user views to modify topic coverage of a test questionnaire;
13 provide appropriate user views for online queries regarding statistics of cells and item analysis;
14 provide appropriate user views to generate reports on statistics of cells and item analysis; and
15 all files imported by the users shall not infect the system and its database

4.2 The Non-Functional User Requirements

The non-functional requirements of the eTest Builder depict the overall performance of the system. Relatedly, these requirements also drive the technical architecture of the system [19]. The series of RAD workshops have clarified issues that in turn resulted to the identification of constraints comprising the non-functional user requirements of the system. The system must be:

1 accessible by all of the registered users with no specific operating system (OS) platform and computer machine assignment;
2 accessible by all registered users with no required specific brand of browser; and
3 able to allow interface with any printer available across the users’ environment.

4.3 Conceptual Framework

Fig. 2 shows the road map used for the development of the eTest Builder intended for a nursing education like the SUCN.

Figure 2: The Conceptual Framework of an eTest Builder
The development of an eTest Builder took off applying the assessment theories and concepts, the curriculum and various syllabi initiated at the SUCN, and a software quality business model. Using a defined scope for its development inclusive of the three main modules, the developed system was governed by a software quality dynamics that emphasizes functionality, usability, reliability, and efficiency. As an eLearning tool equipped with particular features, the automated test questionnaire builder as the end product – the eTest Builder – is expected to boost productivity among the faculty members at the SUCN.

5 THE SYSTEM

Developed observing high involvement of users as mandated by RAD, the eTest Builder is a system reflecting the SUCN’s needs and wants with respect to their assessment practice and technology integration. The user-interfaces, establish the link between the system and the user to be able to communicate, are screen layout charting characteristics of capabilities of the system providing clear instructions for the user. A login mechanism is a scheme to establish stringent security to maintain the test bank integrity and the basis to establish command responsibility. The following items show the selected screenshots of the developed eTest Builder:

5.1 The System Administration Module

The System Administration Module covers the management of the curriculum used along with individual syllabus of each subject, the users (faculty members), the team-teach groups of the particular semester, and the housekeeping of the test bank. Fig. 3 is the screen shot that elucidates features under this module.

5.2 The Test Bank Module

The Test Bank Module is responsible for the management and calibration of the test bank along with features on the management of test questionnaires including creation of coverage for a test questionnaire, test questionnaire generation, and processing of the item analysis. Fig. 4 shows the screenshot implementing the features of this module.
5.3 The Report Management Module

The Report Management Module handles the queries about the test bank’s contents and to generate required reports to aide SUCN’s decision-making endeavour on their assessment practice. The module provides an interface for the user to establish criteria in the information generation. All results of the online queries and the generated reports — graphical or tabular, can be transmitted to an available printer for the generation of hard copies. Fig. 5 elucidates the user-interface of this module.

5.4 The Sample Reports

The implementation of the interactive querying and data visualization results to the generation of relevant information that will assist SUCN in streamlining their assessment practices. The following generated reports are just a few samples of what the eTest Builder can generate:

5.4.1 A sample generated Test Questionnaire

Fig. 6 displays a sample generated test questionnaire implementing clinical case presentation with respective items. The last page of the generated test questionnaire is the answer key of the test.
5.4.2 A sample of an Online Query Report

Fig. 7 shows a sample screenshot of the result of an online query for the statistics of a particular test questionnaire.
5.4.3 A sample of a Data Visualization Report

Fig. 8 shows a data visualization generated report from an online query summarizing all the Item Analyses of a particular nursing subject.

Figure 8: A Data Visualization Sample Report Summarizing All Item Analyses of Subject’s Test

6 SYSTEM EVALUATION

The eTest Builder Users’ Training conducted last April 10, 2015 necessitated the SUCN faculty members to actually use the system over a period of five days. Survey questionnaires adapted from the ISO/IEC 9126-1 served as research instrument were distributed to the attendees of the training to ascertain the software level quality of the completed eTest Builder. On April 16, the researcher retrieved the duly filled-up survey tool from the SUCN secretary. The retrieved survey questionnaires were validated to determine invalid responses. Of the 27 faculty members for the school year 2014-2015, 20, or 74 per cent, who voluntarily participated in the survey, were the official respondents. The survey obtained a 100 per cent yield.

6.1 The Quality Assessment of the Three Modules

Table 1, depicts the internal consistency, using Cronbach alpha reliability coefficient, of the items attributing to the constructs that calibrated the quality assessment of the three modules of the system. The Cronbach alpha reliability coefficients for the quality assessment of the three modules ranged from 0.77 to 0.93 revealing an acceptable reliability coefficient with reference to Nunnally’s 0.60 or greater as acceptable value [20]. Hence, these results proved satisfactory attributing to their internal consistency.

Table 1. Cronbach Alpha Reliability (Validation Statistics) and Descriptive Statistics of the Module Software Quality Assessment Levels of the eTest Builder

<table>
<thead>
<tr>
<th>Scales</th>
<th>Items</th>
<th>Valid Cases</th>
<th>Alpha Reliability</th>
<th>Mean</th>
<th>sd</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Test Bank Module</td>
<td>3</td>
<td>20</td>
<td>0.77</td>
<td>4.57</td>
<td>0.59</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>B. Test Questionnaire Generation Module</td>
<td>2</td>
<td>20</td>
<td>0.83</td>
<td>4.80</td>
<td>0.41</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>C. Statistical Report Generator</td>
<td>2</td>
<td>20</td>
<td>0.03</td>
<td>4.78</td>
<td>0.43</td>
<td>Extremely Agree</td>
</tr>
</tbody>
</table>
The Test Bank Module variable construct got an assessment level of “Extremely Agree” verbal interpretation, from an aggregate mean of 4.57. Similarly, the variable constructs on the Test Questionnaire Generation and Statistical Report Generator Modules both garnered an “Extremely Agree” verbal interpretation, from aggregate means of 4.80 and 4.78 respectively.

### 6.2 The Software Quality Assessment Level

Table 2 showcases the treatment of the Cronbach alpha reliability coefficient to check on the internal consistency of the items attributing to the constructs that calibrated the software quality assessment levels the eTest Builder. The Cronbach alpha reliability coefficients for the five (Functionality to Security) constructs attributing to the software quality of the system ranged from 0.72 to 0.88 revealing an acceptable reliability coefficient proving a satisfactory internal consistency.

<table>
<thead>
<tr>
<th>Scales</th>
<th>Items</th>
<th>Valid Cases</th>
<th>Alpha Reliability</th>
<th>Mean</th>
<th>sd</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Functionality</td>
<td>4</td>
<td>20</td>
<td>0.83</td>
<td>4.61</td>
<td>0.55</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>B. Reliability</td>
<td>4</td>
<td>20</td>
<td>0.72</td>
<td>4.66</td>
<td>0.57</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>C. Usability</td>
<td>5</td>
<td>20</td>
<td>0.88</td>
<td>4.59</td>
<td>0.64</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>D. Efficiency</td>
<td>2</td>
<td>20</td>
<td>0.77</td>
<td>4.70</td>
<td>0.52</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>E. Security</td>
<td>2</td>
<td>20</td>
<td>0.77</td>
<td>4.53</td>
<td>0.60</td>
<td>Extremely Agree</td>
</tr>
<tr>
<td>F. Overall Assessment as a business solution</td>
<td>1</td>
<td>20</td>
<td>-</td>
<td>4.80</td>
<td>0.41</td>
<td>Extremely Agree</td>
</tr>
</tbody>
</table>

The variable constructs Functionality, Reliability, Usability, Efficiency, and Security have aggregate means of 4.61, 4.66, 4.59, 4.70, and 4.53 respectively fall under one verbal interpretation of “Extremely Agree.” The Overall Assessment of the eTest Builder garnered an aggregate mean of 4.80 with an “Extremely Agree” verbal interpretation equivalent.

### 6.3 Relationships between overall and software quality assessments

Table 3 depicts the results of a chi-square computation to establish significant relationships between the overall assessment of the eTest Builder as a business solution to each of the constructs of the software quality. The cross-tabulation of all constructs of software quality to the overall assessment scale revealed all significant statistical association. These indicate that with high overall assessment level of the eTest Builder as a business solution can be translated to high software quality assessment as depicted in Table 2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pearson Chi-Square</th>
<th>Spearman Correlation</th>
<th>p-value</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Functionality</td>
<td>15.313</td>
<td>0.698</td>
<td>0.009</td>
<td>Significant</td>
</tr>
<tr>
<td>B. Reliability</td>
<td>15.833</td>
<td>0.686</td>
<td>0.007</td>
<td>Significant</td>
</tr>
<tr>
<td>C. Usability</td>
<td>16.975</td>
<td>0.712</td>
<td>0.010</td>
<td>Significant</td>
</tr>
<tr>
<td>D. Efficiency</td>
<td>11.667</td>
<td>0.741</td>
<td>0.009</td>
<td>Significant</td>
</tr>
<tr>
<td>E. Security</td>
<td>8.333</td>
<td>0.644</td>
<td>0.016</td>
<td>Significant</td>
</tr>
</tbody>
</table>

* alpha = 0.05
7 SUMMARY, CONCLUSION AND RECOMMENDATIONS

7.1 Summary

The eTest Builder, an application system for assessment in academia, is a by-product of theories of assessment in the field of education – manual and electronic in nature. This application system further used a conceptual framework for an eTest Builder as development roadmap, taking into consideration input elements such as cases, stems, item keys, testing regulations, test construction rules, Nursing subjects offered at the SUCN, the counterpart topic outlines of each Nursing subject, and the variable constructs of software quality as defined by the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC) 9126-1.

An electronic test questionnaire builder is the main output of this study, that resulted in an automated assessment application system that is fast, highly-dependable, user-friendly, and efficient.

The systems development life cycle (SDLC) of the eTest Builder sailed on using a method called rapid application development (RAD), a software engineering discipline meant to reduce its process time without jeopardizing the quality of the manufactured system. Its major feature is the reduction of data modelling approaches that normally would increase documentation time. Instead, RAD employs a great deal of user-involvement, and an appropriate tool like prototyping techniques to capture the requirement of the users, and shorten the feedback intervals.

After all the modules were completed, the system was presented to varied ICT groups known to the researcher for their technical comments on its use. After the successful conduct of the Users’ Training, the SUCN faculty members were given five days to use the eTest Builder. Likewise, they were also administered a survey that asked them to evaluate the eTest Builder software quality as defined the ISO/IEC 9126-1.

The survey conducted among the SUCN faculty obtained a 100 per cent yield, with overwhelming positive results. Overall, the SUCN found the eTest Builder as an excellent business solution for the existing issues on their assessment practice. This general statement is based on the aggregate mean of 4.80 which equivalent to a verbal description of “Extremely Agree” for the overall assessment variable construct.

7.2 Conclusion

This study found that an eTest Builder would address problems on assessment practice like adherence to valid recycling and compliance to the test construction rules. The Valid Recycling, Item Analysis, and the integration of the Table of Specifications (TOS) in the compliance to the Bloom’s Taxonomy are innovative features of the eTest Builder that automates a tedious but highly ideal practice in Educational Assessment. Likewise, the eTest Builder will encourage the use and application of clinical cases or scenarios in making test questionnaires allowing students to be exposed to Nursing Board-like type of exams.

**Robust and secure Test Bank.** When the eTest Builder was tested, it was found that the power of the test bank can be attributed to its inherent design which develop with the end in mind: implementing valid recycling, item analysis, thematic coverage of the test questionnaire anchored in the context of the TOS, and test questionnaire-generator conforming to the rules and guidelines in educational assessment.

**Timely, relevant, accurate Test Questionnaire-Generator.** The timeliness of the test questionnaires generated by the eTest Builder can be attributed to the systems design that implemented the introduction of clinical cases or scenarios. These Nursing cases are a source of information from where test items can be drawn. The use of clinical cases or scenarios is encouraged in the tenets of learning science to promote critical thinking, as it follows students to use their higher order thinking.

The valid recycling mechanism also supports timeliness of the test questionnaires, thereby, the possibility of making use of the same item question but being able to change the key in the question which automatically changes the correct answer among the options.

The accuracy of the generated test questionnaires can be attributed to the algorithm used when the eTest Builder was developed. The business solution used conforms to the basic tenets in educational assessment management, and applies the concept of simple-to-complex sequence of test items as one good attribute of a testing instrument.
Relevant Statistical Report for informed decision-making. The design was a product of the series of RAD workshops during the requirements analysis and design phases; hence, they conformed to what SUCN believed were requirements needed as tools for decision-making.

Functional, reliable, usable, efficient, secure eTest Builder. Results of the survey show that the four variable constructs contributory to the definition of software quality as described by ISO/IEC 9126-1, and the additional variable construct for security all garnered a resounding assessment level of “Extremely Agree” verbal descriptions. This is based on the computed aggregate mean of 4.61 for functionality, 4.66 for reliability, 4.59 for usability, while 4.70 and 4.53 for efficiency and security, respectively.

This implies that the eTest Builder’s software quality anchored on the ISO/IEC 9126-1 description is very functional, approves-reliable, predominantly usable, exceptionally efficient, and highly-secure.

7.3 Recommendations

This study aimed to draw a business solution to the issues on educational assessment practice that hounded one nursing school for a decade. In so doing, this study also aimed to bridge the gap between the mandate of Silliman University through its strategic plan of integrating eLearning in the curriculum development, and the availability of technology. Results of this study point to recommend the following: incorporate the adoption and actual use of the eTest Builder as a College policy:

1. for the College leadership to document every year level’s experience in the use of the eTest Builder with regards to software quality such as functionality, reliability, usability, efficiency, and security;

2. for the College leadership to establish benchmark policy and endeavours with other academic institutions offering the Nursing degree that may have earlier adopted a similar automated generator of test questionnaires;

3. for future innovations of the eTest Builder, other features can be considered like a) support to mobile computing without jeopardizing security, b) level-up to computer adaptive testing (CAT), and c) enhancement of business intelligence (BI) that includes automating validity of items through the use of natural language processing (NLP), enterprise reporting to cover metric management, balanced scoreboards, ad hoc analyses, and data mining.

REFERENCES


