DESIGN AND DEVELOPMENT OF A SELF-ASSESSMENT TOOL AND INVESTIGATING ITS EFFECTIVENESS FOR E-LEARNING

Manisha Domun [manisha.domun@gmail.com], Lifelong Learning Cluster,
Goonesh K Bahadur [g.bahadur@uom.ac.mu], Virtual Centre for Innovative Learning Technologies, University of Mauritius, Reduit, Mauritius [http://vcampus.uom.ac.mu/vcilt/index.php]

Abstract

One of the most effective tools in e-learning is the Self-Assessment Tool (SAT) and research has shown that students need to accurately assess their own performance thus improving their learning. The study involved the design and development of a self-assessment tool based on the Revised Blooms taxonomy Framework. As a second step in investigating the effectiveness of the SAT, 1st year student of the BSC Educational Technology program from the VCILT, University of Mauritius were used as testing sample. At this stage the SAT was provided to only half of the sample who were randomly chosen and placed into a treatment group. The remaining half (*Control Group*) had the normal conditions on the E-learning platform. A semester exam was devised and administered to the whole sample to find out if there was a difference between the scores of both groups. Lastly a feedback form was given to only the treatment group to find out their views on the SAT. The results indicated a significant difference in scores between the treatment and the control groups when the Student's Independent T-test was used. Group A percentage of passes were higher compared to Group B. Failures were recorded for both groups with an increased rate of failure for Group B compared to Group A. Moreover, most of the respondents' feedbacks suggested that SAT was a useful guide with helpful feedbacks. The findings concluded that SAT was viewed more as a revision tool that allowed them to assess their own learning.

Keywords: Self-Assessment Tool, Intellectual Property and Copyright in Educational Processes, e-learning, effective tools, Revised Blooms Taxonomy

Introduction

E-learning provides the opportunity for students to engage into self-directed learning, that is, students take the initiative and the responsibility for their own learning. In this context, the use of Self-Assessment Tools (SAT) has emerged as a great tool of interest for the educational community. Students tend to look for self-assessment tools to assist them with their course materials (Marinagi, 2011). Brahmawong (1991) stated that the most important factors facilitating effective self-learning are active participation, immediate feedback, success experience, and gradual approximation. Schulz (2005) highlighted the crucial role of self-assessments as being valuable for learners in online learning environment, specifying the purpose of design as a tool to:

- Measure the degree to which content is meeting learning objectives;
- Provide an opportunity for immediate application of content;
- Test learners on new knowledge;
- Provide meaningful feedback to learners;

- Track learner progress;
- Focus Learning;
- Make learning fun, interactive or flexible;
- Assist in limiting or providing access to new information.

Schulz (2005) stated that multiple choices, true and false, fill in the blanks, drag and drops are examples of effective online self-assessments. Though simple in design, these assessments are used to test knowledge, add interactivity, and provide immediate feedback to learners. Kostons et al. (2012) pointed out that for self-regulated learning to be effective, students need to accurately assess their own performance on a learning task, and to select an appropriate new learning task in response to that self-assessment. Following the design of both objective and subjective items, it is widely accepted that online learning requires interactivity to improve learners' skills and deliver results.

Park (1996) added that generating effective feedback and adapting to learner characteristics have been understood as important goals for educational technology since at least the early 1960s. Vella (2002) suggested that formative feedback is to be provided at scheduled points for students to improve their learning. Crump (2005) as well emphasized that self-assessment refers to the involvement of learners in making judgment about their own learning, particularly about their achievement and the outcomes of their learning. Chang (2009) further stipulated that when feedback provided is in response to an individual learner's specific learning status expressed by his or her homework; this allows learning to take place.

Most of the available self-assessment tools are not adapted to the Mauritian context. Moreover little has been written about the effectiveness of self-assessment tools in contributing to learners' performance in specific modules pertaining to the Mauritian context as well as students' perception of a Self-Assessment Tool in learning. Therefore this study contributes to research by addressing this context gap.

The theoretical framework

The chosen framework for this study was the Revised Blooms Taxonomy. The RBT consisted of two dimensions namely the Cognitive and Knowledge dimensions which were used as backbone to design all the interactivities of the SAT.

Together these two processes are represented in a two dimensional table as shown in Figure 1. A cell of the table as shown in the diagram is where the knowledge and Cognitive process intersect.

Figure 1. The Taxonomy table

Moreover Anderson and Krathwohl (2001) identified nineteen specific cognitive processes that further clarify the scope of the six categories. Figure 3 is a visual representation of the six levels of the cognitive dimension. The verbs in the first three columns describe the lower order skills whereas the last three columns represent verbs describing the higher order skills.
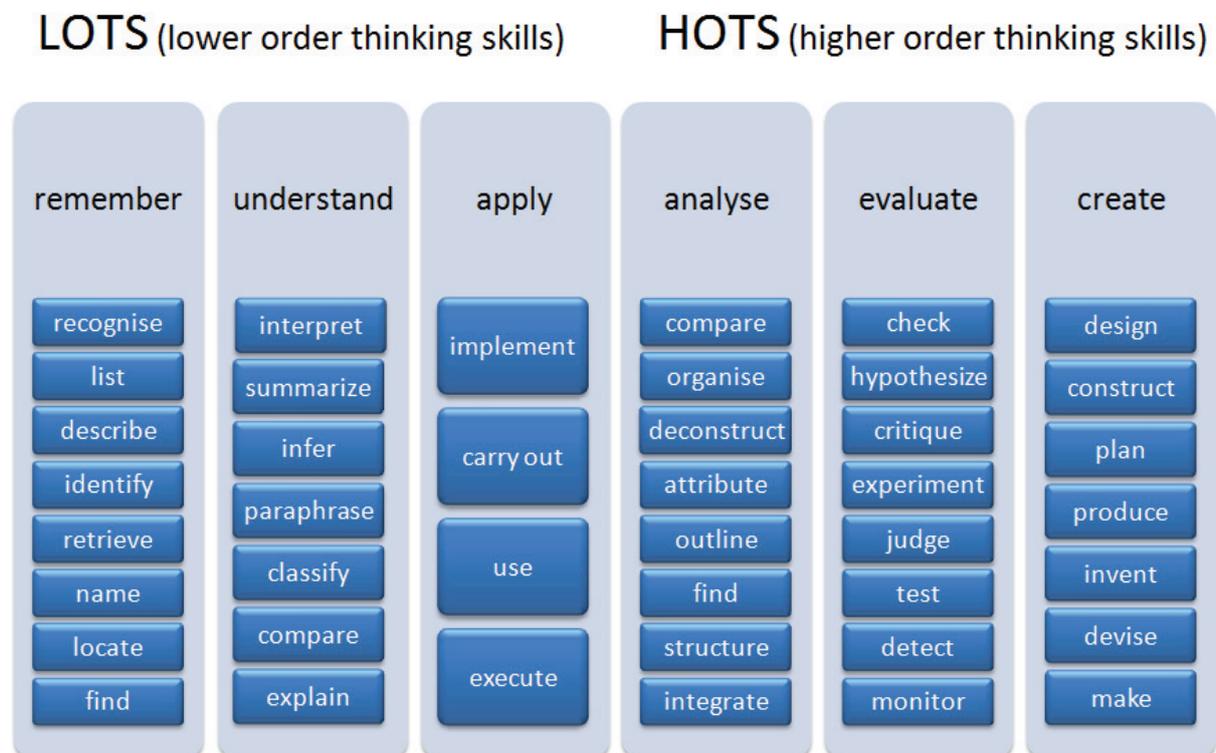
Figure 2. The Revised Blooms Taxonomy verbs

Figure 2 outlined the verbs that defined how the cognitive dimension of RBT starts from lower to higher order skills. Therefore when designing the SAT, emphasis was laid on using questions that started from lower to higher order skills for each interactivity. Detailed application of the RBT is outlined in the Methodology section

Aims and objectives

The experimental study determined whether the intervention of using SAT had the intended effect on the students learning of a particular module namely Intellectual Property and Copyright in Educational Processes which is a core module in the Programme of BSC (Hons) Educational and Instructional Technologies offered at the University of Mauritius (http://ilearn.uom.ac.mu/lms/course/view.php?id=265).

The Experimental Approach was divided into three different phases:

1. Design and development of a Self-Assessment Tool (SAT).
2. A Performance Test was given whereby only half of the sample had access to SAT followed by a Post Test (Exams) where the whole sample had access to SAT. Additionally, marks were compared to marks obtained from previous years, for better reliability and validity of the results.
3. Views of the students were obtained to investigate the effectiveness of SAT towards their performance.

Methodology

Sample population

The sample used in this study consisted of undergraduate students enrolled for the BSc (Hons) Educational and Instructional Technology offered by the Virtual Centre of Innovative Learning Technologies (VCILT) at the University of Mauritius. All students enrolled on this programme were educators with a minimum teaching experience of 5 years. A sample population of 40 students was selected. They were randomly chosen and placed into two distinctive groups:

4. **The treatment group**: The treatment group denoted by Group A consisted of 20 participants randomly chosen from the undergraduate students enrolled for the BSc (Hons) Educational and Instructional Technology at the VCILT. They were placed in a test group created on the VCILT *ilearn* platform and provided with the same learning resources as the control group but with an additional tool, the Self-Assessment Tool (SAT) two weeks prior to the written test.
5. **A control group**: The control group denoted by Group B consisted of the remaining 20 students from the sample of undergraduates students. They were given the normal conditions with the learning resources but were not provided with the SAT. Both groups would take the same written test.

Data analysis

The data was analyzed both quantitatively and qualitatively. For this study an independent t-test was used as statistical method for the analysis of results obtained for the performance test. Descriptive statistics was used to compare the results obtained for the performance test with that of the Exams results. The SPSS software was used to analyze the quantitative as well as the responses obtained from the qualitative data.

Statistical analysis

Both pre-test and post-test were performed in strict examination condition with two invigilators assigned.

To assume normality a box and plot diagram was used.

The Box Plot

A Box Plot represents the visual assessment of the normality (symmetry) of the data as well as the equal variance assumption. Figure 3 represented the different sections in the Box Plot.
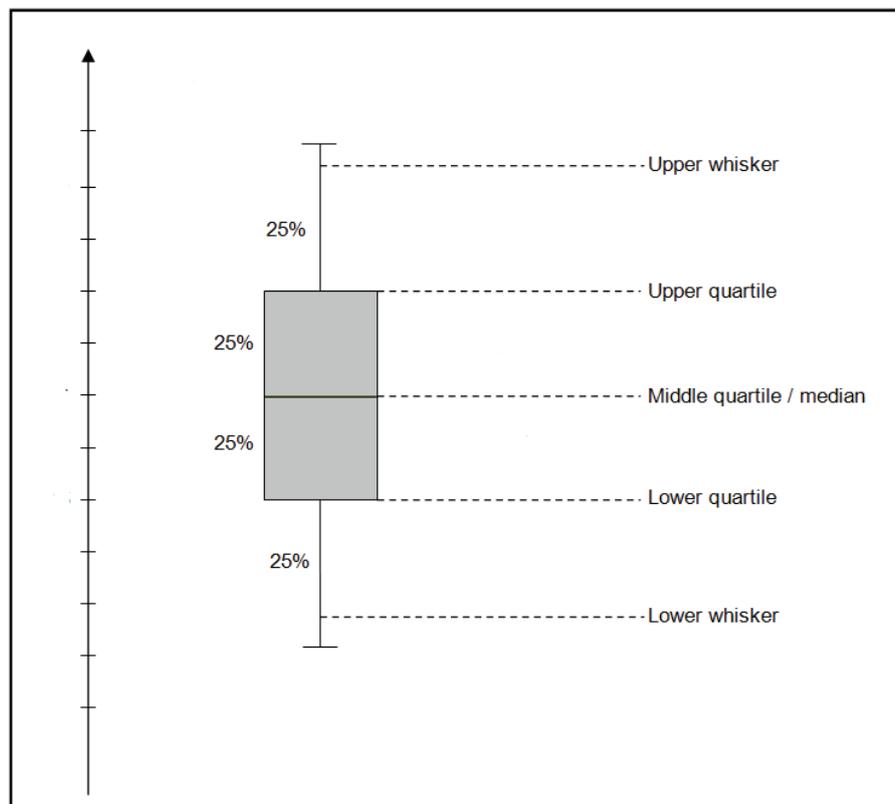

Figure 3. Box and Whisker Labels

The Box plot was a useful way to show overall patterns of response for both group A and B as well as visualize the range, distributional characteristics of both groups' scores. Observation method was used to relate pertinent findings prior to the test. The results of both assessments were performed using descriptive statistics. The **Levene's Test** was carried out to determine whether the population sample variances for the two groups were equal and whether the two groups were significantly different in their means.

A two-sample Student's t-test assuming equal variances using a pooled estimate of the variance was performed to analyze the scores obtained from the performance test.

Is there a statistically significant difference between the scores obtained by the students who used SAT and those who did not used SAT?

It is not necessarily that the causal force is the use of SAT. The hypotheses for this test were:

- H_o: $\mu 1 = \mu 2$ versus H_a: $\mu 1 \neq \mu 2$ where $\mu 1$ is group A mean score and $\mu 2$ is group B mean score.
- The **Null Hypothesis (H_o)** test that there is no significant difference between the two group scores
- The **Research Hypothesis (H_a):** There is significant difference between the two group scores.

For this study, a confidence level of **95%** was used.

- If p value < 0.05 , then the Null Hypothesis of no difference in scores will be rejected and H_a will be considered

If the p value > 0.05, then the Null hypothesis will be accepted and H_a will be rejected.

Feedback analysis – Survey

Qualitative data analysis was intended to obtain students' experience with the SAT and relate its effectiveness in triggering learning using both open-ended and closed-ended questions.

Designing the SAT

The Revised Blooms Taxonomy (RBT) consists of two dimensions namely the Cognitive and Knowledge dimensions which were used to design all the interactivities in the SAT. Screenshots have been included to support the detailed assumptions of the cognitive and the knowledge dimension of the RBT.

The SAT was developed with Articulate Quizmaker software and which consisted of different interactions such as "True and False", "Fill in the blank", "Matching", "Multiple Choice" (One and Multiple correct answers), "Essay questions" and "Time graded quiz".

In designing the questions for the SAT, tweaking and prioritizing objectives were crucial to prioritize the content of the module "Intellectual Property and Copyright in Educational processes".

Priority 1: What learners must know?

Priority 2: What learners should know?

These were answered by the tutor during the informal interview where the following priorities were highlighted as to what learners must and should know from the module "Intellectual Property and Copyright in Educational processes" namely:

1. They must know the difference between Intellectual Property and Intellectual Property rights
2. They must understand the different types of rights, specific terminologies related to Copyright issues.
3. They must know the difference between trademark and patent.
4. They should apply the relevant type of protection that is most appropriate for any given scenario.

5. Learners should be able to highlight the possibilities, and limitations of fair use concept in educational context (both in classroom and online learning scenarios)

6. They should have a global understanding of the rights as well as the Conventions and National Legislation laws pertaining to Copyright and its use in education.

The interactivities chosen and implemented for the SAT aimed at assisting learners in attaining both lower and higher order skills in the learning of the outlined contents. The following cases give an overview of the SAT interactivities used and how they were adapted to the Revised Bloom Taxonomy (RBT) lower and higher order skills to ensure retention and learning of the theory based content.

Implementation of guidelines

To maximize use and effectiveness of the SAT, a guideline was provided. A link was additionally provided on the E-learning platform to guide learners on how to access and use the SAT. Both written guidelines and demonstration through a video with voice-over features were provided to ensure optimum access to the different interactions in the SAT. The Guideline tool was developed using the **Camstasia 8 software.**

Test cases

Test cases were used to show how the SAT was designed to meet up the requirement outlined in the RBT.

Test case 1: Catering for factual knowledge

This was achieved through 3 types of interactivities namely:

- True-false item
- Fill in the blank
- Matching activity

These activities consisted of lower order thinking devised to help learners for memorization and simple recognition Figure 4 shows an example of a true/false item intended to test level 1 of the cognitive dimension of the RBT namely 'to remember facts'. The design process also included some questions which were levelled to meet higher order skills to promote students reasoning skills. Learners were requested to recall what intellectual property is and select "true or false" to validate the answer.
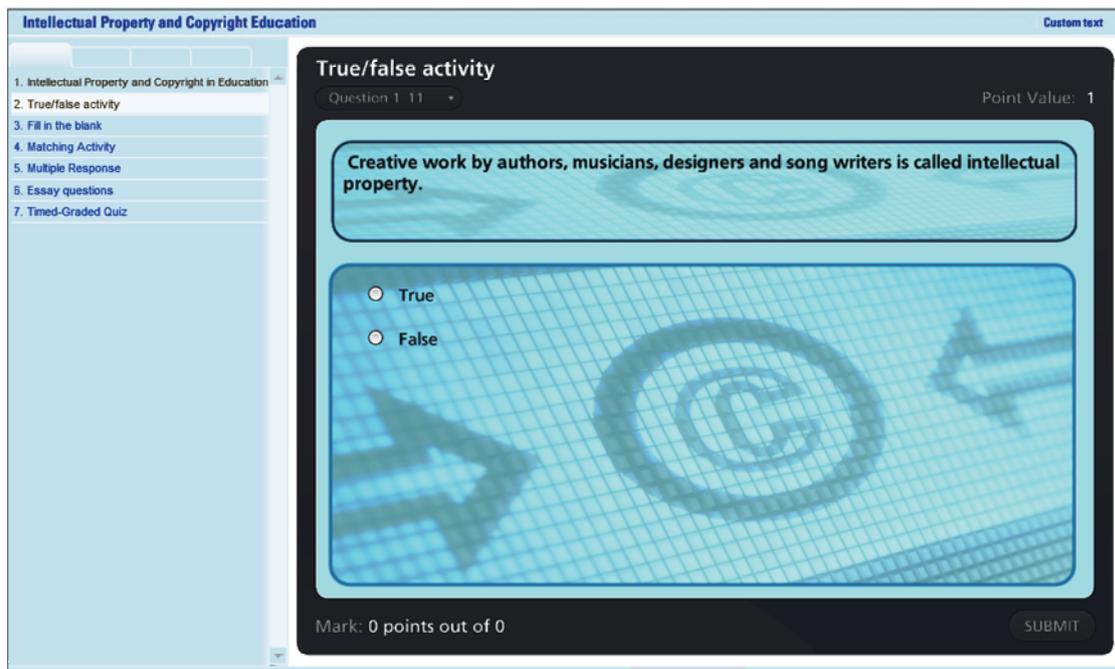
Figure 4. True-False item: Remember for factual knowledge

Figure 5 shows the "fill in the blanks" interactivity activity. The activity promoted reasoning skills providing learners with the ability to recall key terms pertaining to a description where they needed to input the missing words.
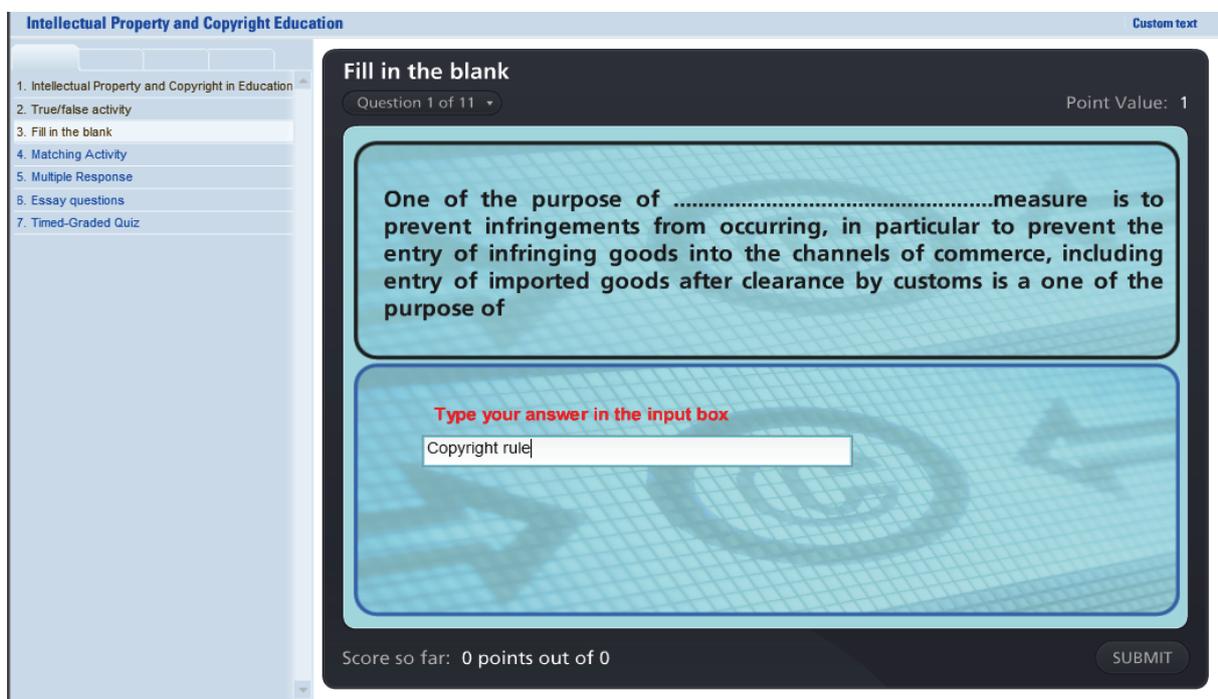

Figure 5. Fill in the blanks: Recall definitions

In Figure 6, learners were required to match the definition by choosing from the drop down list. This item tested ability to recall conceptual knowledge and match definitions.
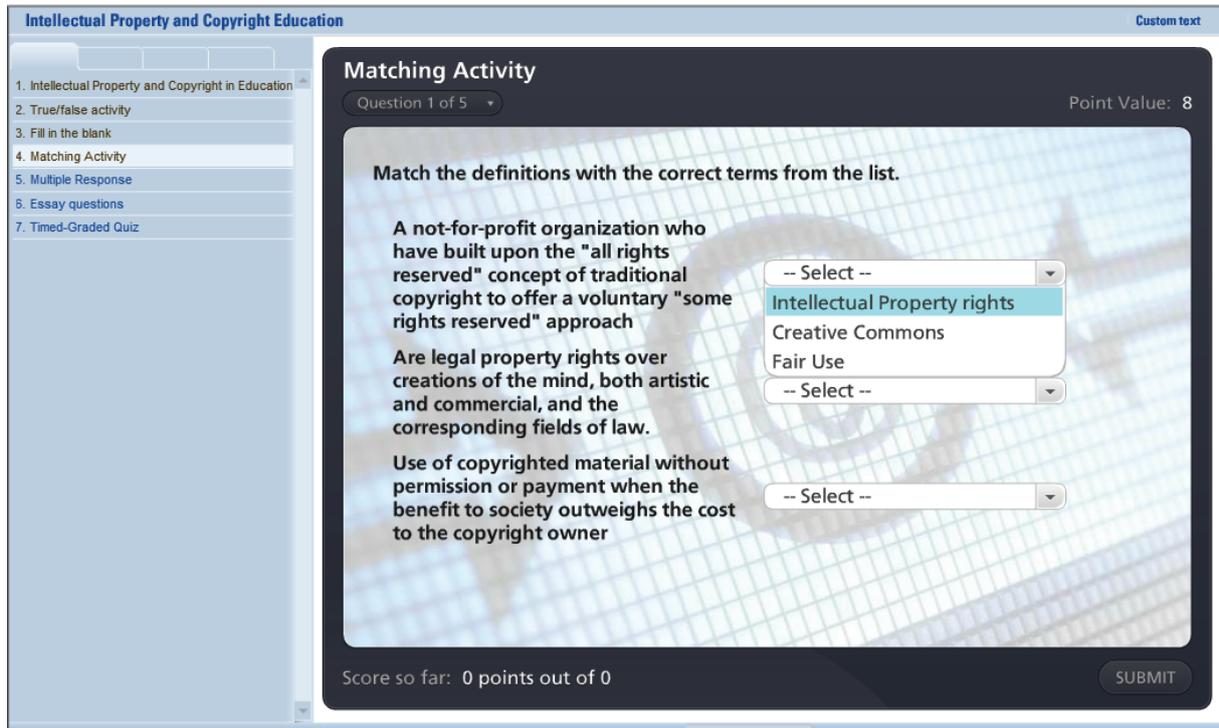
Figure 6. Matching activity to memorize key terms and associated definitions

Test case 2: Interactivities planed at understanding level

Case 2 showed some of the interactivities catering for level II of the RBT that is the understanding level to assist learners in assimilating definitions, along with novel applications of theories and principles related to the "Intellectual property and Copyright" module. At the understanding level learners required interpretation skills and the true/ false activity was levelled to make the learners derive the meaning.

Test case 3: True-false levelled at understanding level

The True-false item, example shown in Figure 7, tested conceptual knowledge through exemplifying. Learners were required to click on either true or false to validate the answer. The question defined a problem situation where the learner had to understand the implications of copyright infringement and select either true or false. The true-false questions were extended from simple recall to understanding.
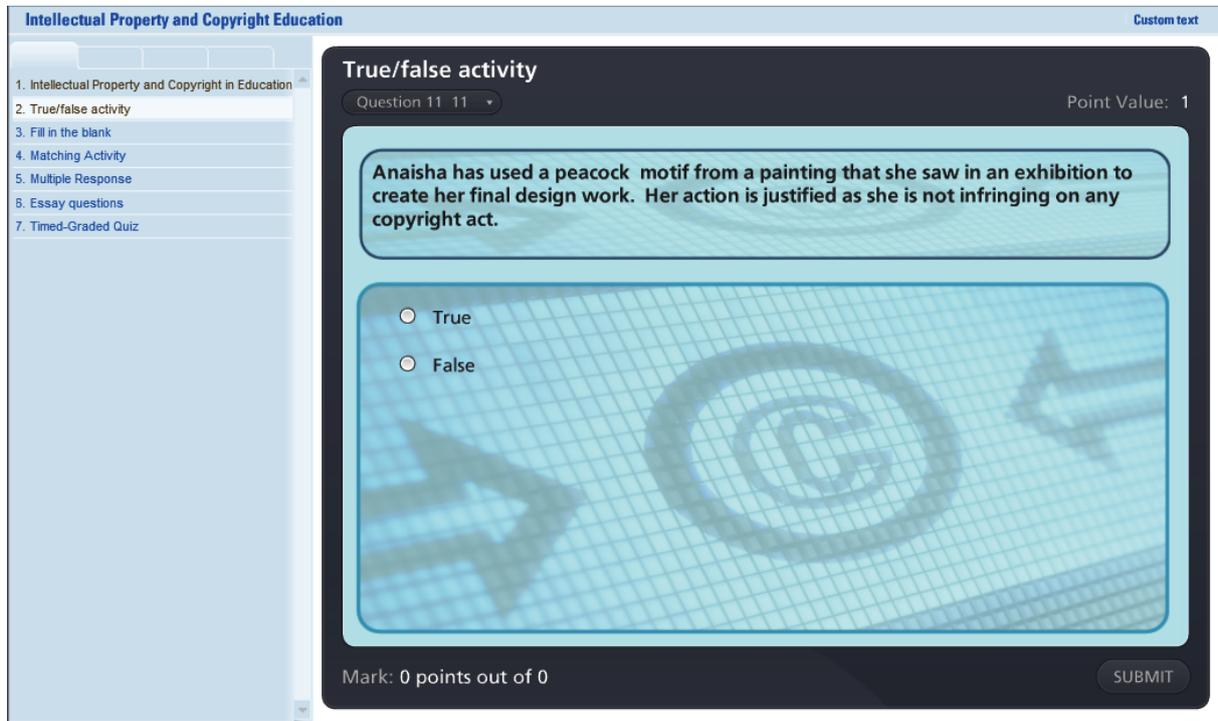
Figure 7. True-false activities levelled to RBT Understanding level evaluating conceptual knowledge

Test case 4: Testing conceptual knowledge with time graded quiz

The time graded interaction included questions that tested the different levels of knowledge of the RBT. The time graded quiz intended to test all levels from factual to meta-cognitive level of knowledge and additionally specifying the relative point value and the approximate time limit.

Figure 8 shows an example of a contextual question requiring learners to apply and adapt their knowledge to a specific situation. Learners were presented with a context whereby they were required to answer based on their knowledge. The aim was to promote complex thinking.
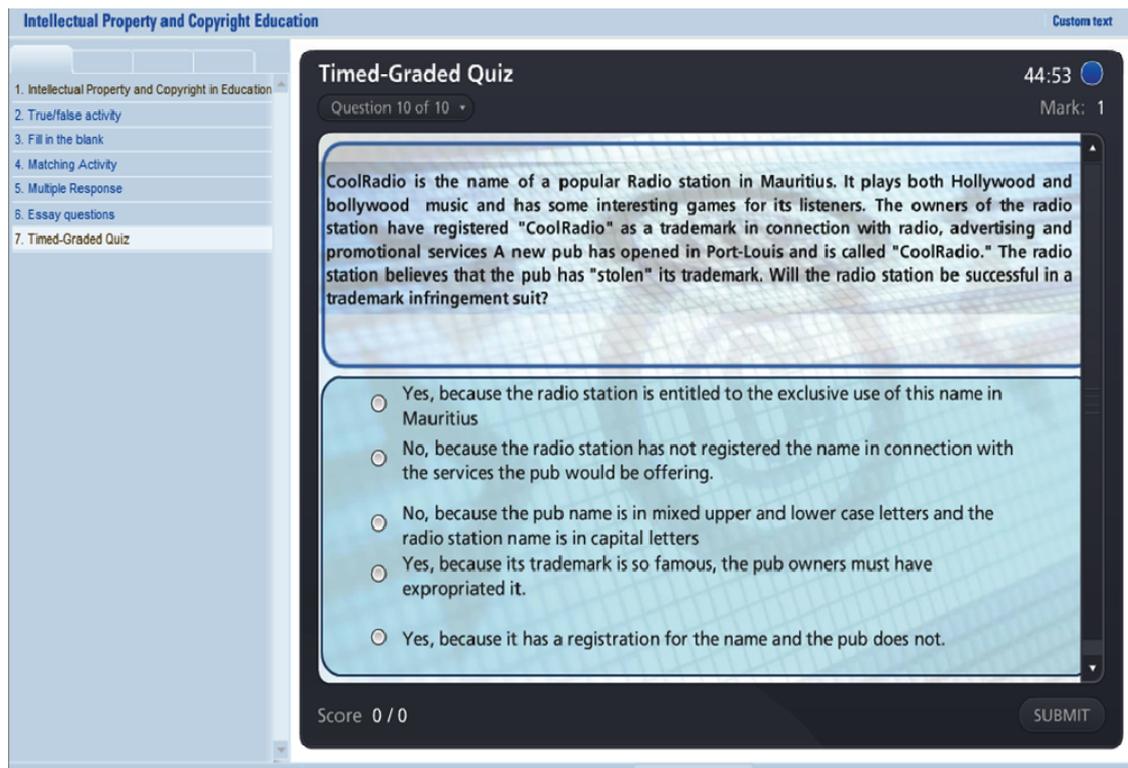
Figure 8. Time-graded Activity testing conceptual knowledge

Test case 5: Testing application level using time-graded quiz

The application level tested the ability of the students to apply the different concepts pertaining to intellectual property and copyright to a given situation. This was implemented through the time graded quiz, where learners were given scenarios and they had to choose the action to be taken from a list of choices. Different knowledge processes were tested. The following screenshots show the interactivities in the given SAT to test learners' ability to apply concepts.

Application level objectives were tested through the "Time-graded quiz", "Multiple Response items" to enable learners to use knowledge of procedures, rules as well as principles or theories and also to apply their knowledge to novel reality based situation.

The "time graded quiz" was designed to match different level of cognitive skills. In Figure 9 students had to apply their knowledge to a given scenario provided and chose the correct answer.
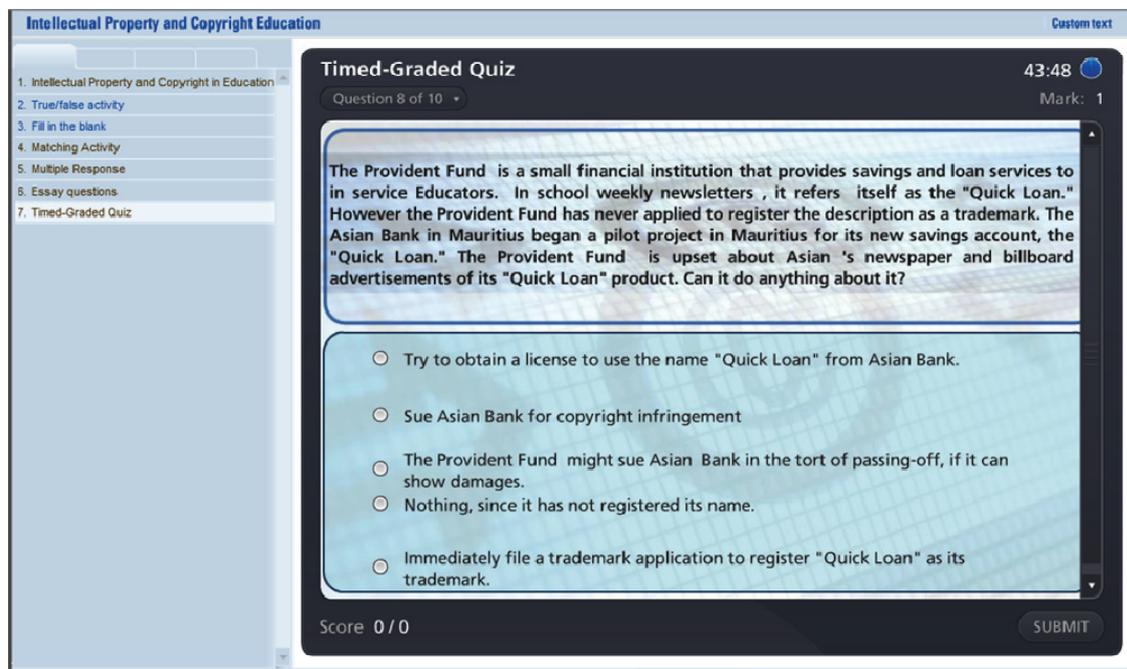
Figure 9. Testing Application level using Time-graded quiz

Test case 6: Testing analysing level using multiple response

The analyzing process referred to the ability for students to break a concept down into parts and describe their relation to the whole.

Analysis level objectives were tested to enable learners to:

- Distinguish between critical and non-critical information;
- Breakdown behaviours or communication;
- Interpret relationships between things;
- Examine information components to answer the question.

This was implemented through the "Multiple Response item" where the questions set provided a mix of item levels as shown in Figure 10. Learners were requested to examine the information in the question and distinguish the critical from the non-critical information.
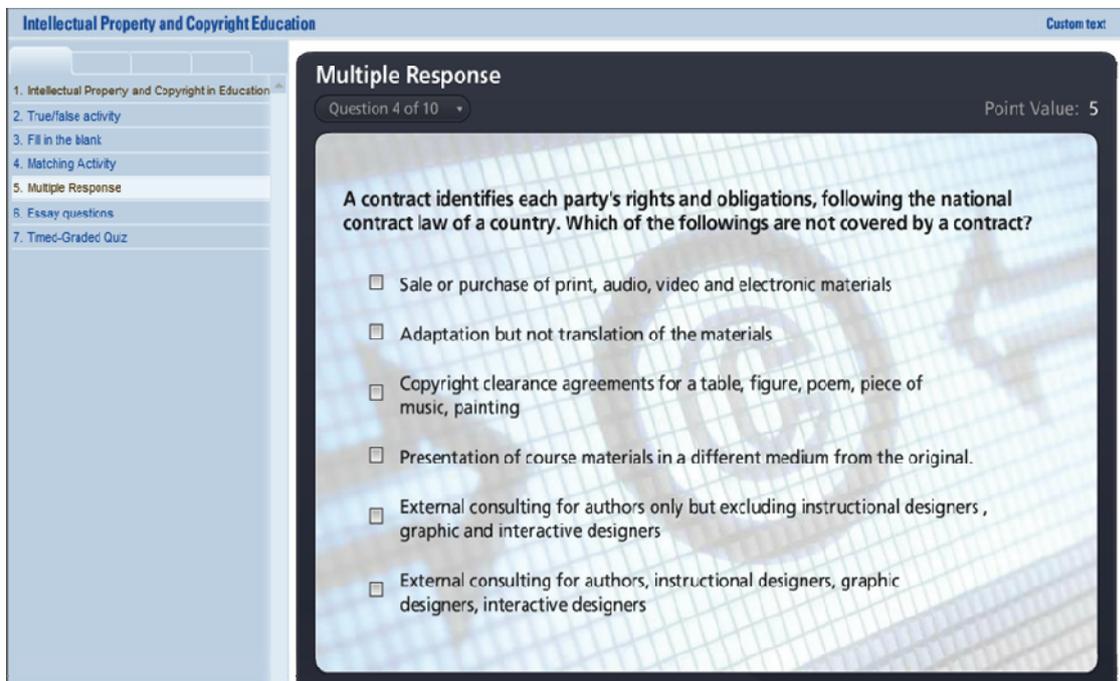
Figure 10. Testing Analyzing level using Multiple responses

Test case 7: Testing evaluation and creativity aspect

The Evaluation and Creativity aspect were tested through Essay Types questions. The objectives were to allow learners make judgments about the aspects pertaining to Copyright and Intellectual Property in Education Processes. Figure 11 showed implementation of an Essay Type question where learners had to provide judgment and prioritize use of concepts.
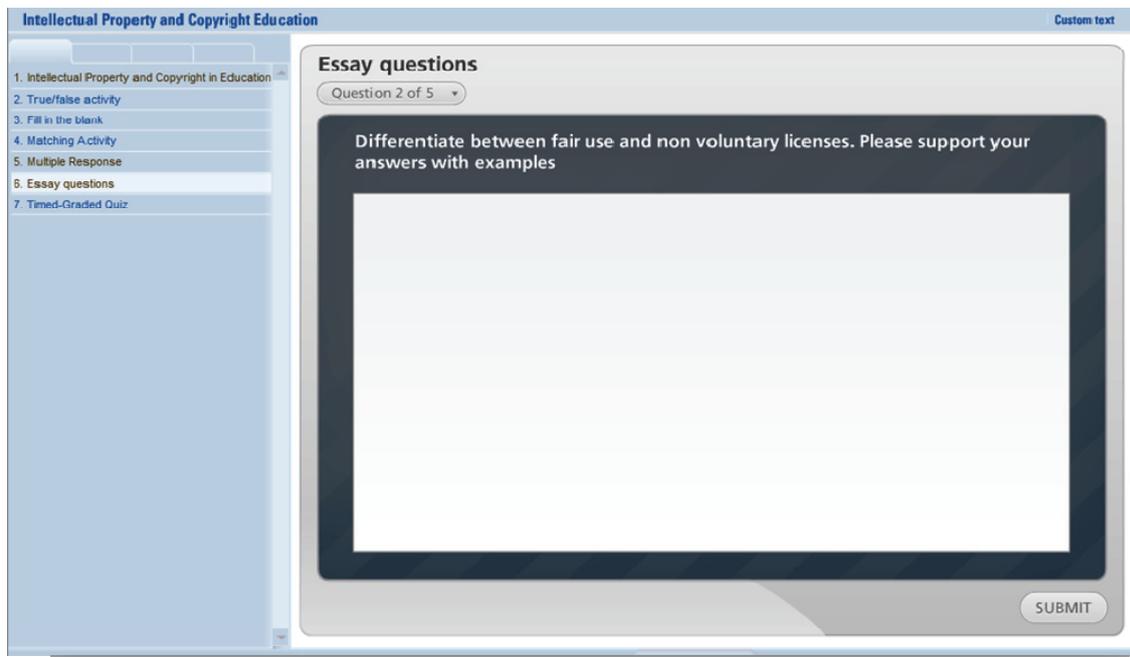
Figure 11. Testing Evaluation level using Essay questions

Test case 8: Feedback mechanisms

One essential element that was deemed crucial for the design of the SAT was the feedback timing. The aim was for learners to get feedback while they were still mindful of the learning target and for them to get feedback while there was still time for them to act upon.

Feedback was implemented in the SAT through:

- Immediate feedback through explanatory feedback to questions of fact (Figure 12);
- Immediate responses to learners misconceptions (Figure 13);
- Summative feedback at the end of each interaction (Figure 14).

Figure 12. Explanatory feedback

Figure 13. Immediate feedback
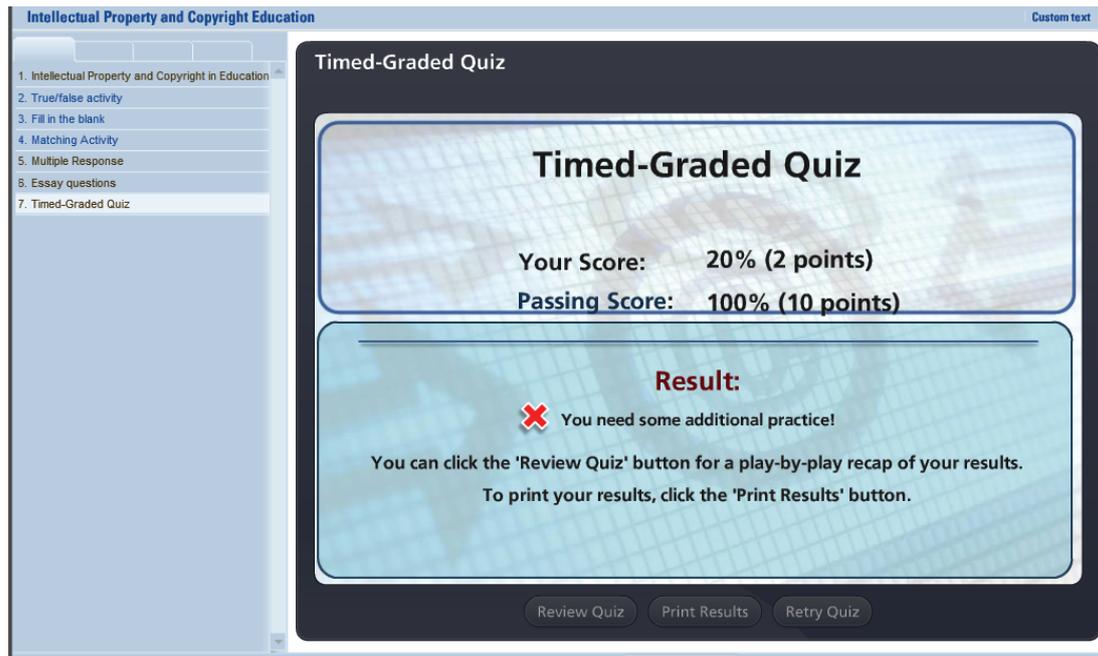
Figure 14. Summative Feedback

Test case 9: User friendliness

The SAT consisted of a left hand side menu allowing learners to navigate and choose their activities. Useful features implemented in the SAT for easy navigation and enhanced learning experiences included:

- The ability for students to leave any activities anytime they wanted to. This was helpful to "because we don't want to force users to click on every element, especially those learners that learned segments of the content".
- The SAT was designed for multiple trials enabling students to try the interactivities at their own convenience.

Result and analysis

The Levene's test result

The "Independent Samples Test" provided two statistical tests. In the left two columns of numbers (Table 1) is the Levene's Test for Equality of Variances for the assumption. As per Table 1 the "Levene's Test," yielded a p-value (Sig.) of p =0. 565 where p > 0.05. This implied an insignificant F-test thus accepting the sub null hypothesis of no difference between the variances. Since the variances of Group A and Group B were equal, the first column of the independent sample test was used, with equal variances assumed.

Table 1: T-test Result for Performance test

Pre-test	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% confidence interval of the difference	
								Lower	Upper
Equal variances assumed	*.338*	*.565*	*2.234*	*38*	*.031*	*4.850*	*2.171*	*.455*	*9.245*
Equal variances not assumed			2.234	37.627	.031	4.850	2.171	.454	9.246

P-value result

The p-value for the equal variances t –test referring to Table 1 was p=0.031. With the p value less than 0.05, the decision taken was that there was evidence of "a significant difference between the two groups scores (Group A and Group B)" thereby rejecting the Null hypothesis (H_o).

Confidence level

A 95% confidence interval on the difference between the two population (Group A and Group B) means using a Student's t-distribution with 38 degrees of freedom was (0.45, 9.24), which confirmed that there was a statistically significant difference in the population means of Group A and Group B.

The 95% confidence levels for the difference in means were given by the last columns of the table 2. All plausible values of the mean difference of the two groups (Group A and Group B) lied between the range (0.45 and 9.24) whereby t (38) =4.85.

The following statistics also confirmed the result of significant difference:

- The interval range did not include '0'.
- The conclusion at 0.05 level of significance level was rejected for the mean value

Table 2: Comparative study of marks obtained for Performance test and Exams marks

Performance Test over 50 Marks			Exams Over 100 Marks
	Treatment Group A (SAT)	**Control Group B**	**Total Marks (SAT access to both groups)**
Total students examined (number)	20	20	40
Mean score	29.15	24.3	
%Percentage passed	85%	75%	80%
%Percentage failed	15%	25%	20%

From the first two column of the comparative in Table 2**,** Group A (SAT treatment) results were compared with Group B (No SAT treatment) the following observations were made:

There is evidence within the observed setting that group A students scored better on average t(38)=2.33, p=0.03 (Table 1). From the observation made during the performance test, students from group A faced less difficulty when answering the questions. Group A had 85% pass and 15% failed compared to Group B who had 75% pass and 25% fail. The mean scores obtained for group A (M=29.15, SD=7.199, N=20) were not significantly different from the scores for group

B (M =24.30, SD=6.51). In other words equal variances were assumed, with an equal spread of scores. Both groups had failures but the rate of failure was greater for group B.

Post-test results

When the samples of 40 students were examined at the end of the semester, 80% pass rate was recorded. The marks were computed over 100 marks with pass mark of 40.

One significant observation, however, was that more failures were recorded in group A (5 failed) than Group B (3 failed). Overall a percentage pass rate of 80% was recorded which was very encouraging. For the performance test, Group A performed better than Group B and this was partly due to the fact that Group A were given the SAT two weeks prior to the test. However when both groups were given SAT for the exams, Group B performed better compared to Group A. Pertinent observations were made and are as follows:

- Unlike the first set of results, group B performed much better compared to Group A. Equal number of failures were recorded for both pre-test and post-test (8 out of 40 students failed the exams)
- One student from Group B (B17) scored 24/50 without the SAT in the pre-test but when exposed to the SAT scored 88/100.

This was very encouraging and affirmative of SAT positive influence on improving performance

Comparison with past results

The results of the post-test were compared to previous cohorts of the past two years. The biases which may result in selection of comparison groups were limited as students from previous cohorts had the same profiles as the current participants. All of the participants were either in-service primary or secondary educators.

One crucial aspect noteworthy were that both previous cohorts were given the normal learning conditions available on the "Intellectual and Copyright Module" currently available on the ilearn platform.

Table 3: Details the percentage passed and failure rate for the past three cohorts

Cohort	Percentage passed (%)	Percentage failed (%)
2010 – 2012	72%	28%
2011 – 2013	65%	35%
2012 – 2014	80%	20%

The students' performances on these tests however were not sufficient to evaluate the educational quality. Popham (1999) implied that standardized achievement tests should not be used to judge the quality of education. He additionally stated that "The overarching reason those students' scores on these tests do not provide an accurate index of educational quotient."

Feedback results from survey

The findings from the feedbacks obtained after using the SAT showed that 70% of the respondents understood the meaning of SAT while 30% of the sample could not provide any definition. Moreover they reported that SAT piloted on the platform was their first experience shown. Among the definitions given by the 70% of respondents for SAT were:

- A tool used for self-learning
- A tool to allow learners to have meaningful feedbacks.
- Test level of understanding on certain topics
- A tool to evaluate how far teaching and learning has taken place
- A tool allowing a person to correct his own mistakes through feedback.
- A tool used for self assessment of one's own capacity and getting feedback
- SAT allows learners to test their understanding and learning. It gives instant feedback and provides remedial action upon wrong responses.

It was noted that the 6% of students were not able to provide a definition for SAT. The students were queried about their prior experience with SAT in their learning. 95% of the students never used SAT for their learning purpose and one student who used SAT specified that it was through quiz and multiple choices provided on the internet.

Relevancy

70% found the SAT relevant to their purpose as well as found the content consistent with the objectives of the course. 30% of respondents termed the SAT as very relevant and appropriate to them as learners. None of the respondents found SAT irrelevant to their content. 65% of respondents found that the SAT interactivities were very good at assessing a wide range of learning outcomes. This was a very positive indication of students attaining the objectives set when using the SAT interactivities to assess their learning. This complied with Baume (2009) study on writing and using good learning outcomes where he stated that:

> *"Good learning outcomes are scaffolding on which students build their studies and their learning. Outcomes should confine learning beyond ensuring the module is coherent and contributes to the program".*

One key point noted was that none of the respondents selected 'fair', 'poor', or 'very poor' as response. This was indeed very encouraging regarding the effectiveness of the SAT interactivities.

Effectiveness of the SAT prior to RBT framework

Students had different perceptions regarding the interactivities implemented for the SAT. 45% of respondents found SAT very detailed, transparent and justifiable. 50 % rated it as good and 5% rated it as 'fair'. The Revised Blooms Taxonomy (RBT) was effective in attaining educational objectives of learning range from surface level processing with the True-false activity, Matching Activity as well as the Fill in the Blanks interactivities. It was very encouraging to note that 80% of respondents found these interactivities useful. Feedbacks included:

- Time graded quiz helped them to allocate their time in answering specific questions but they complained of insufficient time to complete the interactivities
- The Time Graded quiz facilitated recall of factual knowledge, promoted remembering, understanding as well as aided learners to apply their conceptual knowledge. However

one common complaint among the students was that they could not cover all the questions in the time set.

- The Multiple Response was very helpful and targeted all the cognitive dimensions of the RBT. However students complained of the lack of immediate feedback provided and preferred single answer multiple choice over multiple responses.

Krathworl (2002) outlined RBT as a framework to classify the instructional and learning activities used to achieve the objectives, as well as the assessments employed to determine how well the objectives were mastered by the student. The present study confirmed this statement as students were engaged across different levels of thinking across the continuum from knowledge to deep level processing or critical thinking. Respondents elaborated that the interactivities:

> *"Covered simple to complex knowledge from recall to formulation of responses"*

> *"Had different level of difficulties which helped in the achievement of learning objectives of the module. "*

> *"Enabled me to apply knowledge to new situations and to solve problems by thinking of new solutions"*

Moreover some students speckled some weaknesses pertaining to the design of the questions:

- Some questions were redundant
- Questions were ambiguous
- Some questions were confusing

20% of the respondents termed the multiple response questions as "confusing". Additional comments included:

- Single answer multiple choices were preferred over multiple response questions which were termed as 'confusing.'
- Multiple choice questions should have been more correctly formulated, i.e. the questions should be written using simple English terms.

60% of the respondents found SAT very useful for remembering and recalling key concepts of the module. The remaining 40% were also positive about the SAT as helping to memorize key concepts. None of the respondents denied the fact that SAT was useful for remembering key concepts. Moreover students specified that interactivities such as True /False, The Fill in the blank as well as the Matching activity were more effective to test recall of factual knowledge. Expanded comments used to term these interactivities were:

> *"It was straightforward and concise in terms of simple, clear definitions."*

> *"The Matching activity helped to recall key words and match them to their respective definitions."*

40% of students rated SAT interactivities as excellent in promoting better understanding through contextualized situation relevant to the Mauritian setting. 50% rated them as 'good' and the remaining 10% found SAT interactivities as 'fair". Moreover from the 10% of respondents one student suggested that:

> *"Putting these definitions into context for instance using a case study could have helped to better contextualize the concepts to the Mauritian setting."*

95% of students recognized that the SAT interactivities stimulated higher order thinking skills engaging them into creating, analyzing and evaluate contents. Chang (2009) stated that "Effective learning requires deepened thinking, because learned knowledge might possibly become one's own through the necessary thought processes. If there is no stimulus to stir up one's thinking, one might simply indulge in receiving, but not digesting information. In this way, superficial knowledge is likely to blossom". In this connection, one student highlighted that the SAT interactivities "*made me reflect on each of the individual answers. It was engaging in the sense that it required deep reflection*"

However the student who rated SAT interactivities as 'fair' acknowledged that "*SAT helped me to memorize key concepts mainly*". He viewed the SAT more as a tool to help him understand "*many terms and concepts which were difficult to memorize*"

The RBT was a wise approach to look at the partial alignment of the different interactivities included in the SAT. However blank areas recorded in their performance test called further to reflect on how the interactivities could have been improved and levelled to cater for the higher order skills of the cognitive dimension and engage students in the process of creating factual, conceptual and procedural knowledge.

Time taken

Students had varying experience for the time taken to complete the interactivities. 46% completed the activities in a reasonable (1= fair) amount of time, 39% were rated as (2=good) while 15% rated themselves as (3=very good) at completing the task. It was noted that none of the students termed the time taken for completing the interactivities as '4=poor' or '5=poor'. This result demonstrated that most of the students completed the activities set for the SAT.

Usefulness of hints

None of the participants disliked the hints that were provided as corrective measure for wrong responses given. It was found that 65% of students rated the hints as 'very good'. 30% of respondents rated SAT hints as 'good' with only one student rating it as 'fair'. Nevertheless none of the students selected very poor. The result demonstrated that all the participants were in favour of the way the hints were implemented in the different interactivities of the SAT.

Helpfulness

With respect to the reasons explicating why they preferred the way that feedbacks were given, two themes were identified as a result of the data analysis:

There were several views expounding the theme of Helpfulness. The Helpfulness of the SAT included the following comments from students: very helpful, nice, very fruitful, enriching, a great motivation to learn, meaningful, excellent tool for revision, promoting self learning.

Some discrepancies were as well noted regarding some interactivities which were termed as "Redundant", "Confusing" such as the Multiple Response where students complained of insufficient instructions regarding number of responses to be validated. Additional weaknesses spotted by the students were "*The fill in the blank activity was case sensitive*".

Importance of feedbacks

70% of students found the feedback provided by the some interactivities of the SAT very specific and timely. They rationalized it to that fact that ample feedback was useful, to reinforce their learning of the concepts. However one weakness experienced by some students regarding the feedback mechanism was that not all interactivities had the immediate feedback option. For some interactivities students were compelled to complete the set of questions for a specific interaction first, then they could view their detailed results "*Instant feedback for all questions was requested by most students*".

This was substantiated by Chang (2009) who highlighted the importance of immediate feedback when the participants in his study mentioned that feedbacks helped them a lot in their learning.

The results obtained for this study also coincided with Riffell and Sibley (2003) where they found out that when feedback were given in a programmed standard form, the students felt that frequent and detailed hints (programmed feedback) significantly increase their ability to learn. Riffell and Sibley (2003) further argued, based on the results of a survey study, that immediate feedback was tied to three components useful for effective learning. These included self-motivation, time management, and organization. In this present study, views obtained for feedback pertained to:

- Very helpful in spotting weak areas which necessitated more focus;
- Allowed review and reworking of questions;
- Helped retain difficult concepts within a limited time.

Moreover, based on the statistics obtained most students (70% + 25%) positively supported the notion that timely feedback was useful because:

- The summative feedback assisted them in reviewing their answers and in spotting out their errors.

The findings of this study did not comply with Roediger & Butler (2010) who found that Quizzes with feedback delayed until the end of the assessment gave better results than those with feedback after each question and supported with Riffell & Sibley (2003) who pointed that students' gratitude also alluded to the fact that delayed responses would result in varying levels of student frustration. This was particular observed in this study by one student who rated the feedback as "fair". The student termed the matching activity as well as multiple responses as less useful due to the lack of immediate feedback. He suggested that:

> "*For a question which has been wrongly answered, opportunities for immediate correction could be provided so that the learner gets to correct him/herself before moving on to the next question instead of summative feedback.*"

The student's suggestion likewise complied with a study conducted by El Mansour and Mupinga (2007) whereby 34 online students were surveyed. The researchers found that without the quick feedback, many students would feel lost in cyberspace.

Nonetheless respondents from this study, found the explanatory feedback very constructive. They highlighted that the explanatory feedback engaged them into high-level knowledge building where they engaged themselves in the learning process by discerning the rationale behind the responses provided '*I was less frustrated as the feedback showed me where I went wrong*' Thereby one consideration to be taken from all these studies was that feedback was crucial but specific timing was more important. Formative feedback can be helpful to reduce anxiety.

Tool for revision

Students viewed the SAT mostly as a tool for their revision. Respondents highlighted one important feature which they found particularly efficient in some interactivities namely:

- The students reasoned that questions set were very contextual and matched their subject providing them with a set of questions to practice for the exams. The questions set for SAT guided them to revise and test their understandings of different chapters.
- It was a great tool for revision giving an insight of the potential questions. Additionally concepts were applied in different situations which enable more understanding compared to just reading notes.

The students furthermore reported that SAT consisted of a variety of activities suited for their varying levels hence facilitating understanding of difficult concepts.

Conclusion

This study comprised of developing a Self-Assessment tool for an E-learning module and investigates its effectiveness in affecting students' performance at tertiary level. A sample of 40 undergraduates students were chosen and randomly placed into a treatment and a control group. A pre-test was given as a Performance Test where only the treatment group had access to SAT followed by a post test (Exams) where both treatment and control group got access to SAT. Statistical analysis confirmed a significant difference in scores obtained for both groups. There was a difference between Group A students who scored better than group B for the performance test. It was noted that for the exams however, Group B students performed better. Furthermore failure was recorded for both groups.

However despite statistical evidence of significant difference in the performance test, there were some gaps regarding the effectiveness of SAT in increasing performance. It cannot be concluded that the performance test or exams revealed the students' potential and understanding of the module in a genuine manner. Low achieving scores could possibly be explained by students' different innate abilities to respond correctly to the test items provided. Failure rate recorded could have been due to low abilities of some students who despite receiving SAT failed for both the test and exam.

Nevertheless the present study can be termed as a positive endeavour which has provided evidence that SAT use on e-learning platform definitely contributed to assist students' revision. Students acknowledged that SAT "was a handy tool in guiding revision and that such an initiative could be implemented in other modules to facilitate learning". Moreover they acknowledge that SAT was "Mostly for revision purposes", an "effective tool for revision by encouraging learning by doing". This led to the final conclusion that SAT was mostly viewed as a revision tool with useful instant feedbacks that acted as a roadmap to guide the revision process.

References

1. Anderson, L. (2002). *Curricular Alignment: A Re-examination. Theory into Practice-Revising Bloom's taxonomy.*

2. Anderson, L.W. and Krathwohl, D.R. (eds.) (2001). *A taxonomy for Learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Addison Wesley Longman.

3. Baume, D. (2009). *Writing and using Good learning outcomes.* Leeds Metropolitain University Journal.

4. Brahmawong, C. (1991). *Techniques of Writing Self-Learning Modules.* National Institute for Multi-media Education (NIME), Ministry of Education, Chiba, Japan.

5. BYU Community (2013). *Copyright 101.* Retrieved December 01, 2012, from BYU Copyright Licensing: http://copyright101.byu.edu/

6. Chang, N. (2009). Can Students Improve Learning with their use of an Instructor's Extensive Feedback Assessment Process? In International Journal of Instructional technology and Distance Learning, 6(5), (pp. 49-63).

7. Crump, C. (2005). *Designing meaningful and fair tests and assignments: A handbook for teachers.* Antigua & Barbuda: Printing and Publishing Co.

8. Carless, D.; Joughin, G. and Mok, M.M.C. (2006). Learning–oriented Assessment: Principles and Practice. In *Assessment & Evaluation in Higher Education, 31(4)*, (pp. 395–398).

9. El Mansour, B. and Mupinga, D.M. (2007). Students' positive and negative experiences in hybrid and online classes. In *College Student Journal, 41(1)*, (pp. 242-248).

10. Ghosh, M. (2008). *Creating Assessment Questions in an eLearning Course.* Random Ideas. India.

11. Kostons, D.; Van Gog, T. and Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. In *Learning and Instruction, 22*, (pp. 121-132).

12. Krathwohl, D.R. (2002). *A Revision of Bloom's Taxonomy.* OHIO. The H.W. Wilson Company.

13. Mandernach, B.J. (2003). *Quality True-False Items.* Retrieved 09 October 2012 from Park University Faculty Development Quick Tips.

14. Marinagi, C. (2011). Web-based adaptive self-assessment in Higher Education. In A. Méndez-Vilas (ed.), *Education in a technological world: communicating current and emerging research and technological efforts.*

15. Mayer, R.E. (2001). *Multimedia Learning.* New York: Cambridge University Press.

16. Mayer, R.E. (2002). Rote versus Meaningful Learning. In *Theory into Practice, 41(4)*, (pp. 226-232). doi:10.1207/s15430421tip4104_4a

17. Murayama, K. (2003). Test format and learning strategy use. In *Japanese Journal of Educational Psychology, 51(1)*, (pp. 1–12).

18. Nayak, B.K. (2010). Understanding the relevance of sample size calculation. In *Indian Journal of Ophthalmol, 58(6)*, (pp. 469-470). Accessed on 17 November 2010. http://www.ijo.in/article.asp?issn=0301-4738;year=2010;volume=58;issue=6;spage=469;epage=470;aulast=Nayak

19. Park, O-C. (1996). Adaptive instructional systems. In D. H. Jonassen (ed.), *Handbook of research for educational communications and technology,* (pp. 138–153). New York: Macmillan.

20. Reiner, C.M.; Bothell, T.W.; Sudweeks, R.R. and Wood, B. (2002). *Preparing Effective Essay Questions -A Self Directed guide for Educators.* New Forums Press.

21. Riffell, S.K. and Sibley, D.H. (2003). Learning online: Student perceptions of a hybrid learning format. In *Journal of College Science Teaching, 32(6)*, (pp. 394-399).

22. Gagné, R.M. (2004). *The Principles of Instructional Design.* Cengage Learning.

23. Roediger, H.L., III. and Butler, A. (2010). The critical role of retrieval practice in long-term retention. In *Trends in Cognitive Science, 15*, (pp. 20-27). http://people.duke.edu/~ab259/pubs/Roediger&Butler(2010).pdf

24. Schulz, A. (2005). *Effectively using self assessments in online learning.* 18th Annual conference on Distance Teaching and Learning. Illinois: The Board of Regents of the University of Wisconsin System.

25. The Independent-Samples t-Test. (n.d.). Available at http://oak.ucc.nau.edu/rh232/courses/EPS525/Handouts/Understanding%20the%20Independent%20t%20Test.pdf

26. Vella, J. (2002). *Learning to listen, learning to teach: The power of dialogue in educating adults* (Rev. ed). San Francisco: Jossey-Bass.