Computer-based testing – the ideal tool to assess on the different levels of Bloom’s taxonomy

E de Bruyn
Department for Education Innovation
University of Pretoria
Pretoria, South Africa
erika.debruyn@up.ac.za

E Mostert
Department for Education Innovation
University of Pretoria
Pretoria, South Africa
el-marie.mostert@up.ac.za

A van Schoor
Department of Anatomy
Faculty of Health Sciences, University of Pretoria
Pretoria, South Africa
albert.vanschoor@up.ac.za

Abstract— Computer-based assessment became one of the most common forms of technology enhanced assessment since the 1990s. Bloom’s Taxonomy is widely used as a classification scheme to determine different levels of cognitive competencies. Using CBT for assessment has often been described as a tool to test on the lower levels of Bloom’s taxonomy and that it promotes rote learning because it is very difficult to write items that will test on the higher cognitive levels. A study was done in the Faculty of Health Sciences at the University of Pretoria to obtain the perception of students and lecturers regarding the level and use of computer-based objective assessment. From the study it is clear that it is a suitable assessment method to test on a variety of cognitive levels.

Keywords: Bloom’s taxonomy, objective assessment, MCQs, computer-based testing, computer-based assessment, computer-assisted assessment, e-assessment

I. INTRODUCTION

The use of computers has become increasingly popular in education over the last decade, especially in higher education where the computer plays a vital role in teaching, learning and assessment of students. This paper will focus on assessment on higher cognitive levels using technology, also referred to as e-assessment or computer-based assessment.

Computer-based assessment, also known as computer-based testing (CBT) became one of the most common forms of technology enhanced assessment since the 1990s (Akdemir & Oguz, 2008). CBT as an assessment tool has been implemented by many universities as part of their assessment strategy. CBT is often only associated with multiple choice questions (MCQs) (Marsh et al, 2007; Bull & McKenna, 2004; Cox & Clark, 1998) which are usually knowledge-based. With the use of computer-based testing systems that provide other question types such as multiple response, matching, drag and drop, hot spot and fill in the blank questions, as well as high-resolution graphics, video, sound and animations, it is believed that CBT can be used effectively to test on higher cognitive levels, it all depends on the way the items are constructed. CBT can also be seen as a learning tool for students as it provides immediate feedback (Bull & McKenna, 2004; Peat & Franklin, 2002; Thelwall, 1999) and lecturers can draw different reports and statistics from the electronic systems that provide them with valuable information about the student learning that will assist them to enhance their teaching practice.

Bloom’s Taxonomy is widely used as a classification scheme to determine different levels of cognitive competencies. Bloom (1956) structured this classification scheme on 6 levels:

- **Knowledge**: recall or recognition of specific items
- **Comprehension/Insight**: state a problem in one’s own words
- **Application**: applies what was learned
- **Analysis**: breaking down a concept into its component parts
- **Synthesis**: put parts together to form a whole
- **Evaluation**: make judgements about the value of ideas or materials

According to Cox and Clark (1998) the aim of teaching is to instill a deep level of learning and understanding in students and not only a shallow or surface understanding of concepts. This approach refers to the last three levels of Bloom’s taxonomy (analysis, synthesis and evaluation). Palmer and Devitt (2007) modified Bloom’s taxonomy in three levels where level 1 refers to recall of information on the knowledge level, level 2 refers to comprehension and application while...
level 3 focuses on problem-solving – using knowledge and understanding in new circumstances.

Using CBT for assessment has often been described as a tool to test on the lower levels of Bloom’s taxonomy and that it promotes rote learning (Harmes, 1999) because it is very difficult to write items that will test on the higher cognitive levels as described by Bloom. Harmes (1999) however indicates that a drive to develop items that will be able to do just that (test on higher levels) has emerged due to the unique possibilities that technology provides in contrast with paper-and-pencil tests.

According to Thelwall (2000) there is evidence that CBT and even MCQs can test higher intellectual skills. It is also believed that with careful attention to the quality of MCQs, it can examine many of the same cognitive skills provided by essay tests (Ramesh et al., 2005; Palmer & Devitt, 2007). The challenge is to apply CBT in such a way that it will test on higher levels and promote deep learning.

II. BACKGROUND

The University of Pretoria (UP) has been using computer-based testing since 1991 (Le Roux, 1999) and the use of this type of assessment has increased substantially during the last few years. CBT is used for both formative and summative assessment and therefore receives a lot of attention with regard to developing high quality items. Training courses in designing objective items are presented to lecturers to ensure that the objective assessment items used are of a high standard. The lecturers are also assisted in developing their questions and encouraged to use the valuable statistical analysis (difficulty and discrimination indexes etc.) from the computer-based testing system to improve and enhance the quality of the questions.

The Faculty of Health Sciences at UP, follows an integrated problem-oriented curriculum since 1997, which combines clinical practice with theoretical tuition. The curriculum consists of various blocks where each block presents a total integration of all the different disciplines in medicine with regard to a particular system of the human body (e.g. heart and blood vessels). The “Student Internship Complex” forms the final eighteen months of medical study and consists of pure clinical practice.

Bull & McKenna (2004) stated that not all the outcomes and skills can be assessed by using only one assessment method. This is the case in the faculty where a variety of assessment methods is used to ensure that all the stated outcomes are assessed and students are exposed to different assessment methods in preparation for their professional careers. This implies that CBT should also be integrated with other assessment methods, including written tests, oral examinations, Objective Structured Clinical Examinations (OSCE’s), Objective Structured Practical Examinations (OSPE’s), assignments, portfolios and presentations. Although there are many teaching philosophies employed by UP, the utility model of assessment described by van der Vleuten (1996) is most commonly used. This model states that all assessment: must be valid and reliable, should have educational value for the students, there should be student compliance or acceptability of the content of the assessment, and there should be a positive balance between cost and resources and the educational impact.

CBT is used extensively for both formative and summative objective assessment in these blocks. During the first year of study the level of assessment focuses mostly on knowledge and insight, although students are already challenged to test on the fourth level of Bloom’s taxonomy namely analysis. Assessment from the second year onwards focuses on assessing problem-solving skills i.e. synthesis and evaluation.

The assessment practice has been revised to be more cost- and time effective with regard to the large number of students that need to be assessed. To achieve this while also improving the validity and reliability of the assessment, the assessment plan of several modules was redesigned. Since 2006 the Faculty of Health Sciences started with a total redesign of their objective assessment items in an attempt to enhance deep learning and test on higher cognitive levels.

III. METHODS

The focus of the study was to evaluate the success of redesigning the assessment portfolio for continuous evaluation of large student volumes within the medical curriculum, to include both formative and summative CBT.

This process started in 2007 with the redesigning of existing data banks of questions and the development of new data banks. This was a continuous process and the data used in this study was collected from 2009-2010.

Two blocks were used for this study:
- A first year block (Molecule to Organism) where Clinical Anatomy, Histology, Embryology, Physiology, Genetics and Immunology are integrated
- A second year block (Dissection) where students dissect cadavers and are exposed to the use of wet specimens

In both these blocks the aim was to redesign the assessment component to migrate from time-consuming written examinations to using CBT. The purpose of these tests was to integrate the different disciplines in a horizontal and vertical manner. The redesign also meant liaising with the different departments and lecturers involved in the two blocks. This was a time consuming process and planning and work sessions were scheduled to ensure the participation of all role players. Case studies were presented that required the students to apply their knowledge with regard to specific medical cases. All the different question types that are available in the CBT system were implemented (MCQs, multiple response, matching, hot spots and fill in the blank questions) were structured in such a way that it assessed students on higher levels of Bloom’s taxonomy.

Multimedia was used to enhance these questions and video material and graphics were included. Students were assessed on a weekly basis using questions similar to the examples in Table 1.
<table>
<thead>
<tr>
<th>Bloom’s Taxonomy</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge</strong></td>
<td>Click on 2 examples of cartilage joints.</td>
</tr>
</tbody>
</table>
| **Comprehension**| An injury to the radial nerve (in humeral shaft fracture) will result in the inability to:  
  - flex the elbow  
  - extend the wrist  
  - extend, rotate and abduct the thumb  
  - extend and abduct the thumb  
  - oppose the thumb |
| **Application**   | Click on the vertebral body which is amongst others, essential to be visible when assessing the cervical spine X-ray of a trauma patient. |
| **Analysis**      | Which nerve may be affected in the injury illustrated in the pictures? Please type your answer.  _____ |
| **Synthesis**     | Interpret the image above with the X-ray and then click on the clinical picture below that represents the nerve injury you would expect to see in the patient. |
Evaluation

On the videos, the physiotherapist is treating a seven-year-old boy diagnosed with cerebral motor dysfunction. Based on your observations, which one of the following diagnoses of cerebral motor dysfunction is appropriate for this boy?

- Spastic triplegia
- Spastic hemiplegia
- Spastic monoplegia
- Spastic diplegia

Both the students and lecturers were requested to complete a questionnaire. The questions posed to the students focused on their perception of the use of questions on the different levels of Bloom’s taxonomy. The different levels were explained to them and included in the questionnaire. The number of students that participated in this study consisted of 299 first year and 255 second year students. The student questionnaire consisted of the following statements/questions:

1. Asking questions on higher cognitive levels of Bloom enriched my own learning process.
2. Asking questions on higher cognitive levels of Bloom enrich the content and presentation of Anatomy as a subject.
3. The questions were on such high cognitive levels of Bloom that I could not cope.
4. Were the computer-based test questions on higher cognitive levels than just Knowledge and Insight?
5. Select the average level of Bloom on which you perceived the majority of computer-based test questions.
6. Select the highest level of Bloom on which you perceived the questions.

The results of this questionnaire are presented in Table 2.

The lecturers completed a different questionnaire that focused more on the teaching practice and 12 lecturers completed this questionnaire. The lecturer questionnaire consisted of the following statements:

1. The sampling of questions for the computer-based testing system was adequate for a reliable outcome.
2. The computer-based test questions reflected the learning outcomes of the course.
3. The questions were appropriate for the content being assessed regarding knowledge and skills.
4. The questions encourage deep, active and reflective learning.
5. The computer-based test questions are challenging and it promotes critical thinking.
6. Asking questions on higher cognitive levels of Bloom enriched my own academic development
7. Asking questions on higher cognitive levels of Bloom enrich the content and presentation of my subject.

The results of this questionnaire are presented in Table 3.

IV. RESULTS

Feedback from the students from both first year and second year reflected a positive response concerning their perception of the level on which the questions were asked. Regarding the first question, 381 students felt that by asking questions on higher levels of Bloom’s taxonomy it enriched their learning and 370 students agreed that it enriched the content and presentation of Anatomy as a subject. Only 70 students indicated that the questions were on such high cognitive levels of Bloom’s taxonomy that they could not cope, while 325 students disagreed that the level was too high. From the 554 students, 522 students indicated that the questions tested higher cognitive levels than only knowledge and insight. The results of the last two questions regarding the students’ perception of the average and highest level of Bloom’s taxonomy indicated that they perceived the questions to be distributed between all the different levels.

### TABLE II. STUDENT FEEDBACK

<table>
<thead>
<tr>
<th>1. Asking questions on higher cognitive levels of Bloom enriched my own learning process.</th>
<th>Disagree completely</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Agree completely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st years</td>
<td>4</td>
<td>7</td>
<td>65</td>
<td>172</td>
<td>51</td>
</tr>
<tr>
<td>2nd years</td>
<td>3</td>
<td>5</td>
<td>90</td>
<td>135</td>
<td>23</td>
</tr>
<tr>
<td>Total:</td>
<td>7</td>
<td>12</td>
<td>155</td>
<td>307</td>
<td>74</td>
</tr>
</tbody>
</table>
2. Asking questions on higher cognitive levels of Bloom enrich the content and presentation of Anatomy as a subject.

<table>
<thead>
<tr>
<th></th>
<th>Disagree completely</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Agree completely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st years</td>
<td>2</td>
<td>10</td>
<td>77</td>
<td>178</td>
<td>32</td>
</tr>
<tr>
<td>2nd years</td>
<td>2</td>
<td>11</td>
<td>83</td>
<td>139</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>21</td>
<td>160</td>
<td>317</td>
<td>53</td>
</tr>
</tbody>
</table>

3. The questions were on such high cognitive levels of Bloom that I could not cope.

<table>
<thead>
<tr>
<th></th>
<th>Disagree completely</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Agree completely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st years</td>
<td>47</td>
<td>129</td>
<td>87</td>
<td>26</td>
<td>10</td>
</tr>
<tr>
<td>2nd years</td>
<td>30</td>
<td>119</td>
<td>73</td>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
<td>248</td>
<td>160</td>
<td>54</td>
<td>16</td>
</tr>
</tbody>
</table>

4. Were the computer-based test questions on higher cognitive levels than just Knowledge and Insight?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st years</td>
<td>287</td>
<td>12</td>
</tr>
<tr>
<td>2nd years</td>
<td>235</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>522</td>
<td>33</td>
</tr>
</tbody>
</table>

5. Select the average level of Bloom on which you perceived the majority of computer-based test questions.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st years</td>
<td>55</td>
<td>50</td>
<td>46</td>
<td>66</td>
<td>11</td>
</tr>
<tr>
<td>2nd years</td>
<td>61</td>
<td>27</td>
<td>59</td>
<td>72</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>116</td>
<td>77</td>
<td>105</td>
<td>138</td>
<td>20</td>
</tr>
</tbody>
</table>

6. Select the highest level of Bloom on which you perceived the questions

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st years</td>
<td>52</td>
<td>35</td>
<td>50</td>
<td>61</td>
<td>27</td>
</tr>
<tr>
<td>2nd years</td>
<td>55</td>
<td>23</td>
<td>42</td>
<td>68</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>107</td>
<td>58</td>
<td>92</td>
<td>129</td>
<td>49</td>
</tr>
</tbody>
</table>

Feedback from the lecturers indicated that the majority believed that asking questions on higher cognitive levels was positive in terms of their teaching and assessment experience. Of the 12 lecturers 10 indicated that the questions encouraged deep, active and reflective learning and it enriched the content and presentation of their subject, while 11 lecturers indicated that their own academic development was enriched by this experience.

All the feedback from the lecturers is presented in Table 3 below.

TABLE III. LECTURERS’ FEEDBACK

<table>
<thead>
<tr>
<th>Question</th>
<th>Disagree Completely</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Agree Completely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 The sampling of questions for the computer-based testing system was adequate for a reliable outcome</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>2 The computer-based test questions reflected the learning outcomes of the course</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3 The questions were appropriate for the content being assessed regarding knowledge and skills</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>4 The questions encourage deep, active and reflective learning</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>5 The computer-based test questions are challenging and it promotes critical thinking</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6 Asking questions on higher cognitive levels of Bloom enriched my own academic development</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>7 Asking questions on higher cognitive levels of Bloom enrich the content and presentation of my subject</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
V. DISCUSSION

Students were asked to complete the questionnaire online after their last computer-based test at the end of the block. The different levels of Bloom’s taxonomy were explained to them beforehand and an outline of the levels was presented in the questionnaire where relevant.

Analysis of the student feedback showed that most of the students indicated that the questions set on higher cognitive levels enriched their learning and the use of all the different types of multimedia (sound, video and graphics) helped them to understand the content of the subject matter better. Responses on question 5 indicated that most students perceived the average level of the questions to be on level 4 (analysis), followed by level 1 (knowledge). Students ranked the highest level of the questions in the following order: level 4 (analysis), level 6 (evaluation), level 1 (knowledge), level 3 (application), level 2 (insight) and level 5 (synthesis).

Based on these results it is clear that the majority of students perceived the use of CBT and objective assessment as a positive experience in their learning. Redesigning the assessment and extensive use of multimedia elevated the assessment to a higher level and enhanced assessment as a learning experience.

The majority of lecturers indicated that by structuring their questions to test on higher levels of Bloom’s taxonomy, which has been made possible by using the computer-based system with different question types and flexible features improved their teaching practice. They believed that it encouraged deep learning and adequately tested not only knowledge, but also challenged the students to apply their knowledge. Lecturers however agreed that it is not an easy task to compile these questions and it is time consuming.

VI. CONCLUSION

Objective assessment consists of more than just MCQs and therefore it is a suitable assessment method to test on a variety of cognitive levels. CBT enhances this feature due to the functionalities available e.g. adding multimedia components that would not be possible with paper-based tests.

It is important to take the outcomes into consideration when deciding on a suitable assessment method. The view that objective assessment can only test knowledge-based outcomes has been proven to be not quite true as indicated in the examples in Table 1.

With a continuous growth in student numbers CBT seems to be the ideal tool for both formative and summative assessment due to the fact that large groups of students can be assessed on a regular basis. Immediate feedback is particularly useful when using CBT for formative and self-assessment, which is not always possible with other types of assessment. According to Bull & McKenna (2004) timely and constructive feedback will motivate students to learn more effectively. CBT further lends itself to change an assessment opportunity into an interesting interactive learning experience.

The key to success is the development of appropriate questions and utilizing technology to its fullest capacity.

ACKNOWLEDGMENT

The authors thank Prof JH Meiring previously from the Department of Anatomy, Faculty of Health Sciences, University of Pretoria for his contribution towards this research.

REFERENCES