

# The use of different types of multiple-choice questions in electronic assessment

## U. G. Singh

School of Information Systems and Technology  
University of KwaZulu-Natal  
Durban, South Africa  
e-mail: singhup@ukzn.ac.za

## M. R. de Villiers

School of Computing  
University of South Africa  
Pretoria, South Africa  
e-mail: dvillmr@unisa.ac.za

### Abstract

Open distance learning (ODL) is characterised by increasing numbers of students who are geographically dispersed. Assessment (marking) can be complex --efficient means of assessment are required, but this should not compromise quality. Multiple-choice questions (MCQs) offer unbiased, objective assessment, which can be administered and processed online in e-assessment that provides immediate results.

This study introduces innovative types of questions of the MCQ genre and describes a study, which investigates adoption of e-assessment tools by South African higher-education academics in formative and summative assessment. The research methods employed were a questionnaire survey and interviews. Quantitative and qualitative data was obtained regarding: (1) types of questions used; (2) questions participants believed could effectively assess higher-order thinking skills; (3) types suitable for various levels of study; and (4) the benefits of e- assessment.

The findings of the study have general applicability, but hold high relevance for distance educators.

### INTRODUCTION

This research, mainly set in academic departments and schools at South African tertiary institutions, investigates the usage of online assessment tools in the multiple-choice question (MCQ) genre. In particular, it focuses on the application of novel MCQ formats, types of MCQs that can assess higher-order thinking skills (HOTS), as well as on the levels of study for which the questions are suited.

The time- and location-independent nature of Information and Communication Technology (ICT) can support educators and students alike. Electronic learning (e-learning) incorporates various electronic media including the Internet, intranets, Web-based learning, interactive tutorials, simulations, educational games and online courses (De Villiers 2005; Costagliola and Fuccella 2009). Furthermore, e-learning offers automated assessment, termed e-assessment and used by universities internationally for some portion of their assessment. Tools and systems exist that create, deliver, score, report and analyse both summative and formative assessments, as well as provide customised online services (Khedo 2005). The objective nature of e-assessment, as well as its rapidity and ability to handle high numbers holds advantages for open distance learning.

## **BACKGROUND**

### **Study 1**

In a baseline study in 2009, the researchers studied the extent and nature of usage of electronic assessment tools within Computing-related academic departments and schools at South African tertiary institutions, as well as users' satisfaction with the tools (Singh and De Villiers 2010). Study 1 aimed at establishing a general context for subsequent research by answering two research questions:

1. *What is the current nature and extent of use of electronic assessment in Computing-related departments at South African universities?*

The questionnaire was made available to Computing-related schools and departments. Responses were received from academics at nine tertiary institutions. Usage appeared to be concentrated in six; namely, the University of Cape Town (UCT), the University of the Free State (UFS), the University of South Africa (UNISA), the University of Pretoria (UP), Cape Peninsula University of Technology (CPUT) and Monash University, with more users from Computer Science, where there was a tendency to adopt e-assessment earlier, than from Information Systems and Information Technology. In some cases, the tools and systems had been deployed for more than five years. Although the actual extent of usage was low, it was increasing. There were 36 respondents from nine institutions, of whom 16 were regular users of e-assessment, and 20 were potential users. The systems were used more for formative than for summative assessment. Most usage occurred in situations of high learner numbers, i.e. first-level classes with 100--5000 students. The questions most frequently used were multiple choice and true/false type of questions.

The use of different types of multiple-choice questions in electronic assessment

Deployment of tools is either limited, or supported by institution-wide policies. Although some respondents used e-assessment in an *ad-hoc* way, certain institutions have official procedures, as well as established in-house systems. In such cases there are dedicated laboratories for computer-based testing and administrators to manage sessions. The results of tests and examinations are recorded automatically on learners' academic records and class records.

Despite the small number of users, a variety of tools were adopted by respondents, including Sakai, Vula, CISCO, Blackboard, Moodle, and CompAssess, while others used custom-developed automated marking systems such as UNISA's self-assessment facility on the myUnisa platform, the University of Pretoria's (UP) custom-built Umfundi and Click-UP and tests on various learning management systems (LMSs) at some other institutions (Singh and De Villiers 2010).

## 2. *How satisfied are academic users with their e-assessment tools?*

To determine satisfaction, the questionnaire probed the 16 established users on their perceptions of benefits and disadvantages. The qualitative open-ended responses were mainly in line with the literature. Though few in number, these established users were convinced users. Eighty one per cent of them believe that e-assessment is more effective than traditional forms and gave reasons for their assertion. They pointed out advantages and disadvantages for both educators and learners (Singh and De Villiers 2010).

## **Study 2**

MCQs were investigated further in Study 2 (the present study) through a questionnaire and an interview (face-to-face and telephonic), which aimed to establish:

1. *Which of the various types of questions in the multiple-choice genre do South African academics use?* (The various types are explained in the section called 'TYPES OF E-ASSESSMENT QUESTIONS')
2. What is the relevance of these types for assessing higher-order thinking skills?
3. For which levels of study are the types suited?
4. What benefits are associated with e-assessment?

## **LIMITATIONS**

The study was aimed primarily at academics in Computing Schools and Departments. However, owing to low response levels, non-Computing participants who are active users of the MCQ-genre were also included. The Computing category incorporated the disciplines of Computer Science, Information Systems, Informatics, Information Technology, and Information Sciences. Non-computing users included Educational Technology, Psychology, Economics, Anatomy, Business Management, Mathematics, Family Medicine and Therapeutic Sciences.

## **LITERATURE REVIEW**

### **Overview and purpose of e-assessment**

Multiple terminologies exist; among others, computer-aided assessment (Khare and Lam 2008) and online examinations (Khare and Lam 2008). Our term of preference is electronic assessment (Alton 2009), termed e-assessment (Fielding and Birmingham 2003). Students are presented with sets of online questions, to which they respond, and which are marked electronically. Results are captured and stored in databases and, where appropriate, presented to students.

Well-designed e-assessment tools can enhance diagnostic, formative and summative assessment of learners. Diagnostic assessment ascertains the level of a learner's knowledge before or after learning activities. The purpose of formative assessment is practice and assessment for learning gain. It does not contribute to the final mark (grade), but provides feedback to educators and learners during the course of a module. Summative assessment is formal and structured, used for official grading at specific points during or at the end of a learning programme (Khedo 2005; Khadi 2004; Mc Alpine 2002).

It requires skill, practice and time to write good MCQs and, in particular, MCQs that test higher-order thinking skills (HOTS) (Luckett and Sutherland 2000). The selection of plausible distracters is a major challenge (Mitkov, Ha and Karamanis 2006). Written badly, MCQs can be confusing and demotivating (Alton 2009). Furthermore, the updating of question banks takes time and effort, and may result in inconsistencies in style and quality when different academics contribute to the same bank (Pitenger and Lounsbery 2011).

### **Benefits**

- MCQs offer efficient evaluation in contexts of high student numbers and limited resources (Bani-Ahmad and Audeh 2010).

The use of different types of multiple-choice questions in electronic assessment

- High numbers of tests can be corrected automatically (Costagliola and Fuccella 2009) in a short period of time (Ventouras et al. 2010; Pitenger and Lounsbury 2011).
- Some tools provide immediate, detailed and individualised feedback. Timely feedback can help reduce the gap between actual and desired performance levels (Walker, Topping and Rodrigues 2008).
- In formative assessment, students can repeat the material or progress at their own pace (Luckett and Sutherland 2000).
- MCQs are useful for revision. There are learning benefits when students practice to reinforce their understanding of concepts (Bani-Ahmad and Audeh 2010).
- Automated scoring/marking of assessments in MCQ format is objective and consistent (Ventouras et al. 2010), since judgment is free from bias or inconsistencies of the human marker (Costagliola and Fuccella 2009).
- MCQs, in whatever format, are an efficient tool for measuring learners' achievement, but are best used in combination with other assessment methods (Luckett and Sutherland 2000).

### **Drawbacks**

- MCQs limit the kind of questions that can be asked to elicit understanding (Ventouras et al. 2010). They frequently measure lower-level cognitive skills such as knowledge, memorisation and aptitudes. Concerns exist as to whether they can assess knowledge application and problem solving, or test HOTS, such as synthesis and evaluation (Chiheb, Faizi and Afia 2011).
- Learners are judged on the correctness of their answers and not on how they arrived at them. Moreover, the educator cannot determine whether a topic has been fully understood or not (Ventouras et al. 2010).
- MCQ approaches are criticised for not facilitating active learning, owing to the fact that they do not require learners to explain why they chose a particular option (Pitenger and Lounsbury 2011).
- Guessing allows test-takers to obtain certain scores without actually knowing the material. To counteract this, 'mixed-scoring' and 'negative marking' can be used, deducting marks for incorrect answers. However, this can deter learners from tackling questions on topics where they have intermediate knowledge (Ventouras et al. 2010).

- Data security is an issue if the test-taker's identity cannot be verified (Khedo 2005).
- Facilities must be available where learners can take e-assessments, which is an issue for distance learners (Costagliola and Fuccella 2009). Furthermore, many tools have limited features for disabled users, so accessibility for the physically-challenged must be considered (Maurice and Day 2004).
- For extensive e-assessment, technical and administrative staff must be available to support the infrastructure and procedures (Khedo 2005; Mc Alpine 2002).
- When learners struggle to use a system, they are distracted from their responses. Programmes should be user-friendly and have high usability (Walker, D. J., K. Topping and S. Rodrigues 2008). Furthermore, learners with poor IT skills or who dislike e-assessment may be disadvantaged (Karl et al. 2011).
- Multiple-choice testing seldom allows test-takers to control the sequence and the duration for which questions are displayed (Karl et al. 2011; Walker, D. J., K. Topping and S. Rodrigues 2008).
- It takes time to develop valid test items (Chiheb, Faizi and Afia 2011). Skill, care and practice are required to avoid confusing learners with unanswerable questions or poor alternatives (Alton 2009).

## **TYPES OF E-ASSESSMENT QUESTIONS**

Concerns exist regarding the credibility of e-assessment practices in higher education, particularly regarding its capacity to test higher-order learning. Various objective formats exist in the MCQ genre; some identified in the literature and others provided by participants. Examples are (1) multiple choice, (2) multiple response (3) true/false, (4) true/false with explanation, (5) matching, (6) extended matching, (7) ordering, (8) fill-in-the-blank, and (9) writing or correcting code or text (Costagliola and Fuccella 2009; Chiheb, Faizi and Afia 2011). Some offer enriched testing and can stimulate HOTS (Maurice and Day 2004; Fielding and Birmingham 2003; Khedo 2005; Costagliola and Fuccella 2009; Chiheb, Faizi and Afia 2011). Brief descriptions follow of various question types:

1. Multiple-choice questions (MCQs) and multiple response (MRQs) display lists of answers from which learners choose by selecting one option only for MCQs and multiple options for MRQs.

2. In Extended Matching Items (EMIs) the learner selects the best answer from a list of 10--20 options, which may be used once, more than once, or not at all. There is no need for plausible distracters (Fenderson et al. 1997). EMIs have the benefit of presenting a series of test items on the same theme (Beullens et al. 2002). They take the form of vignettes, where cases are briefly described. EMIs emphasise problem-solving skills and application of knowledge rather than simple recall. They are widely adopted in medical education (Beullens et al. 2002) where, for each item, the students must solve a problem, e.g. by choosing a diagnosis from a list of relevant options. In Matching List questions, learners must specify which items in a list correspond with items in another list. These lists may include text or graphics.
3. True/False questions require learners to decide between two alternatives. The possibility of guessing is a concern, since merely guessing can give an average of 50 per cent, making it easier possible for all learners to pass. These questions are simpler to prepare and quicker to read and answer, so there can be a high number of questions representing the subject matter well. A richer form is True/False with Explanation, where options for the explanation are also provided.
4. In Fill-in-the-Blank questions, learners submit the missing word(s) in a sentence/paragraph or missing items in a table. Hotspot questions require selecting an object or area in a graphic by using the mouse.
5. Drag and Drop Questions test the ability to assign items to the correct category or to arrange parts of a system into a whole by moving icons, images, or textual labels to specific locations on the screen.
6. Simulation Questions present models of systems and require learners to perform a highly interactive task by manipulating values.
7. Short Answer questions require learners to type an answer to a question, typically textual answers to very specific questions.

## **METHODOLOGY**

A questionnaire survey was administered, followed by an interview survey, which gathered quantitative and qualitative data.

### Questionnaire participants

Questionnaire participants were a combination of Study 1 respondents and referrals, both local and international. Questionnaires were distributed by email to 132 potential participants, of whom 92 responded. Sixty four (64) participants were from South Africa and represented 15 institutions. The largest numbers, 13 each, were from UCT and UFS; followed by WITS with 11; UP with 5; and UNISA with 4. Eighteen (18) (19.6%) of the 92 questionnaires were excluded, because they did not provide data.

The research was initially aimed at Computing users, but due to high usage and earlier adoption of e-assessment tools, non-Computing users were also incorporated. The analysis shows composite findings, as well as some results by user type. In the Computing group with 55 participants, 21.1 per cent came from UCT and 18.4 per cent from UFS, with UNISA and UP at 10.5 per cent. Of the 37 non-Computing participants, 30.8 per cent were from WITS, followed by UFS (23.1%) and UCT (19.2%). The Computing users were distributed over 14 institutions, while non-Computing users came from eight. UCT and UFS, with the most respondents overall, had relatively high numbers in both groups.

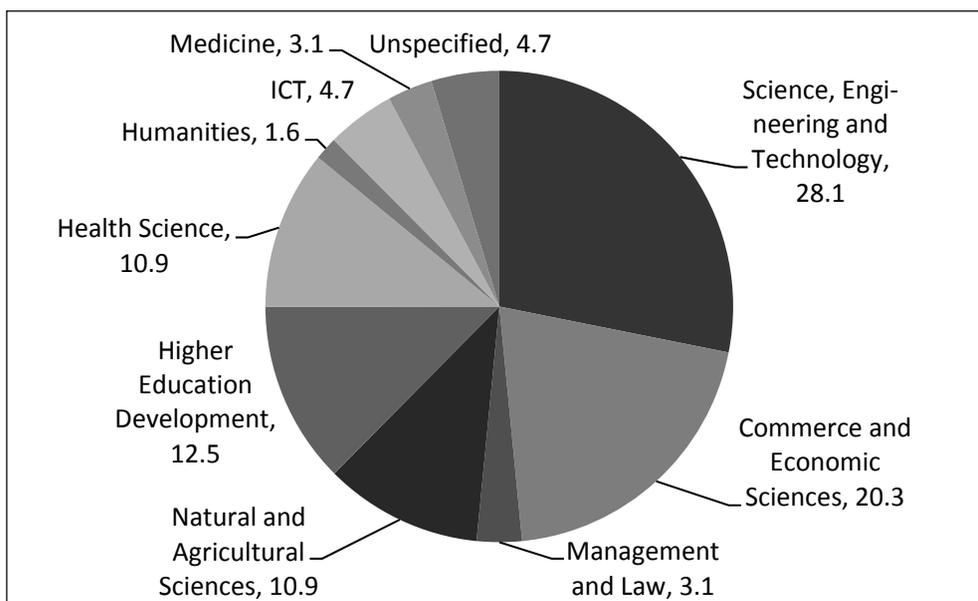


Figure 1: Faculty participation

Figure 1 depicts 11 faculties, with the greatest participation from Faculties of Science, Engineering and Technology (28.1%), Commerce and Economic Sciences (17.2%), and Higher Education (12.5%). Participation from the Health

The use of different types of multiple-choice questions in electronic assessment

Science and Natural as well as Agricultural Sciences was 10.9 per cent each. Health-related departments have traditionally been early adopters of educational technology, partly due to practical work and less time in class situations.

Participants represented 12 schools/departments, with a high number from Schools or Departments of Computer Science (28.1%), Information Systems (14.1%) and Information Technology (9.4 %). There was readiness in the education discipline to adopt technology, with 9.4 per cent from Departments/Schools of Educational Technology (9.4%). Mathematics (7.8%) and Family Medicine (6.3%) also had high participation.

Of the 74 participants whose questionnaires were used, 57.9 per cent were senior lecturers or lecturers. Some participants (34.6% of non-Computing users) had support staff to administer assessments and tests, while certain Computing users -- who were comfortable implementing technology, managed e-assessment personally.

### **Interview participants**

Sixty eight interviews -- some personal and most telephonic -- were conducted with South African academics through referrals from Studies 1 and 2. Twenty three had participated in the Study 2 questionnaire. Of the interviewees, just under half were Computing participants.

Participants represented 11 institutions (including two universities of technology), from eight provinces. Twenty five per cent (25%) were from UCT and 18 per cent from UFS (Qwa Qwa and South campuses), 10% were from UP, 10% from NWU (Vaal and Potchefstroom campuses), and 9% were from WITS. The primary researcher was also referred to four local e-assessment developers, who were interviewed for information on the design requirements of South African academics. Their contributions were particularly valuable in relation to benefits and problems of e-assessment software, as well as the mention of additional functionalities and features used or requested.

## **ANALYSIS AND DISCUSSION OF FINDINGS**

### **Questionnaire**

To understand usage of the various types of questions in the MCQ genre, we addressed the research questions under the subheading, Study 2, in the Background Section.

Sixteen types of questions in the MCQ genre were listed in the questionnaire. The participants were asked to indicate all the e-assessment types they had used over the years and across modules. The most common types adopted by 64 South

African academics, Computing and non-Computing together, were *Multiple Choice: single-response* (40.90%) and *Multiple Choice: multiple-response* (17.44%), followed by *True/false* (14.45%), *Simulations* (10.44%), and the others below 10 per cent. *Text-input, short- answer* had a 14.55 per cent response, but in all instances, these were marked manually.

Participants were also asked to indicate adoption of the various question types in intervals 0 per cent, 1--29 per cent, 30--69 per cent, 70--99 per cent and 100 per cent; e.g. if a participant selects 30--69 per cent for True/False, it indicates that the participant adopts True/False questions 30--69 per cent of the time in their assessments. The adoption patterns of computing and non-computing users, according to this classification, are presented in Table 1.

Table 1: Average adoption of question types -- Comparison of groups

Question Types	Computing (%)	Non-computing (%)
Multiple choice: single response	40.18	41.92
Multiple choice: multiple response	11.99	25.19
True/false	13.61	15.65
True/false with explanation	4.18	4.04
Fill-in-the-Blanks/Completion	9.72	5.37
Simulation	14.68	4.42
Matching Items	5.39	9.19
Extended-matching items	5.39	9.19
Selection/Drop down lists	2.84	4.04
Ranking	5.14	0.58
Diagram/Video Clips	7.14	5.37
Drag-and-Drop	1.62	2.48
Reordering/Rearrangement/Sequencing	1.62	1.15
Categorising	1.62	0.58
Hotspots	1.22	6.71
Text Input (short answer)	15.05	6.09

Chi-square goodness-of-fit tests were performed on this data to see whether, for each question type, there was a significant selection of a specific usage percentage. Despite the indication in Table 1, that computing users seem more willing to adopt a variety of types, there were also significantly more respondents than expected who indicated a 0 per cent usage of the non-standard question types, i.e. no use of types outside direct multiple choice and true/false.

Various cross tabulations were done, one of which, Usage/Faculty indicated a significant relationship in that significantly more than expected respondents from Management Sciences ( $p=.024$ ) were using *Ranking* questions in up to 30 per cent of their assessments.

The second question investigated usefulness of different question types in assessing HOTS. Bloom's taxonomy (Bloom et al. 1956) presents a progression in levels of thinking, starting at the concrete lowest-order Level 1 with *facts*. Thereafter, learners *comprehend* meanings and implications of the facts (Level 2). On Level 3 they *apply* their learning, which helps them to solve problems and transfer knowledge to related situations. In *analysis* (Level 4), learners can classify, categorise, discriminate and detect information, as well as compare and contrast concepts. *Synthesis* (Level 5) involves combining ideas, planning, forming solutions, and creating new information. *Evaluation* on Level 6 requires taking decisions, ranking concepts and making judgments regarding information and situations (Bloom, Mesia and Krathwohl 1964; Passey 2010).

Participants were asked to rate the types in terms of relevancy to HOTS, in four categories: <Not useful>, <Undecided>, <Useful>, <NA/unfamiliar>. Computing and non-Computing users evidenced very similar patterns.

The chi-square goodness-of-fit analysis was done to ascertain which responses were selected significantly more often than the others. Significance is taken for any p-value less than 0.05, indicating 95 per cent confidence. The lower the p-value, the more significant the result.

- <Useful> was chosen by significantly more respondents than expected for *Extended-Matching Items*; *True/False with explanation*, *Diagram/Video Clips*, *Simulation*, *Multiple Choice: single response*; and *Multiple Choice: multiple response*. For these six types, significance was  $p < .0005$ .
- *Ranking* with  $p=.001$ ; *Reordering/Rearrangement/Sequencing* ( $p=.001$ ); and *Categorising* ( $p=.002$ ), were significant to levels indicating high usefulness for assessing HOTS.
- <Not useful> was chosen significantly more often for *True/False* ( $p<.0005$ ); *Matching Items* ( $p=.002$ ); and *Drag-and-drop* ( $p<=.017$ ).

Selection for the response options was not even across options. *Multiple Choice: single response* was the only question type with which all respondents were familiar.

The third question investigated which question types the interviewees felt were useful for the four levels of study.

Table 2: Suitability of questions per level of study (Comp = Computing; Non-Comp = Non-Computing)

Question Types	Year 1 % of Interviewees		Year 2 % of Interviewees		Year 3 % of Interviewees		Year 4 % of Interviewees	
	Comp	Non-Comp	Comp	Non-Comp	Comp	Non-Comp	Comp	Non-Comp
MCQ: Single response	92.1	96.2						
True & false	76.3	80.8						
EMI	73.7	57.7					-	50.0
Selection/Drop down lists	73.7	69.2						
MCQ: Multiple response	81.6	69.2	63.2	69.2			-	57.7
Fill-in-the-blank/ Completion	78.9	69.2						
Hotspots	52.6	-	52.6	-				
Matching items	68.4	-						
Drag-and-drop	68.4	-						
True and false with explanation	60.5	-	65.8	73.1			-	50.0
Diagram/Video clips	60.5	-	57.9	65.4	63.2	65.4	52.6	-
Simulation	68.4	-	65.8	-	63.2	61.4	50.0	61.5
Ranking	68.4	-						
Reordering/ Rearrangement/ Sequencing	68.4	-						
Categorizing	57.9	-	57.9	-				

Table 2 shows that for first-year level, there was a broad use of types, but particularly MCQ (single response) and True/False questions. Computing academics explored every type, especially multiple response, Fill-in-the blank and EMI. On second level, users were selective and made greater use of the more advanced types, while at third level, usage was concentrated on Diagram/Video Clips and Simulation. There was a similar tendency at honours, adoption was greatest on Diagram/Video Clips and MCQ (multiple response).

## **Interview study**

The interview study supplemented the questionnaires by providing further insights into adoption patterns of e-assessment. Of the 68 participants, 52.94 per cent are recent adopters of e-assessment, having had one to two years' experience with it. Only 17.64 per cent had used it for more than five years. According to Costagliola and Fuccella (2009), most online testing modules are part of general purpose learning management systems (LMSs). This was also found to be true in South Africa, with 73.5 per cent of the interviewees using the tool built into their university's LMS.

The study investigated whether participants used a single tool consistently, whether they used different tools for different purposes, or whether they changed their preference. It emerged that 85 per cent (57) of participants had used the same tool for the previous three years. Of the 15 per cent (11) who changed tools, six did so due to institutional changes, while the other five adopted a different tool due to personal dissatisfaction and interest in exploring other options and question types not supported by the tool.

11.8 per cent of the participants indicated that their institutions have a fixed, university-wide policy on the adoption of e-assessment. Academics must conform to the policy and design assessments around it. The other 88.2 per cent were free to use e-assessment as they wished. None of the interviewees were restricted by any school/department-wide policy. All the interviewees had the full support of management in their department/school for adopting e-assessment.

As depicted in Figure 2, e-assessment was used by 70.6 per cent of the academics for formative assessment, which did not contribute to the final mark), while 29.4 per cent used it for summative purposes and 28.6 per cent for both. Marks for these assessments are released immediately by 86.7 per cent of academics, while 13.2 per cent delay release for checking and to prevent copying.

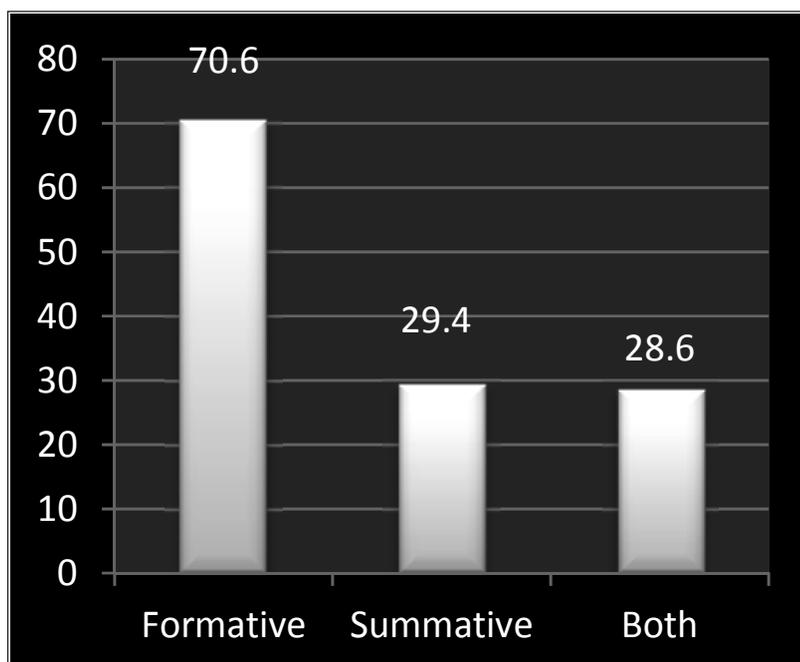


Figure 2: Adoption of e-assessment by type

Table 3 summarises adoption at the various levels. The highest adoption of electronic questions from the MCQ genre occurs at first- and second levels (NQF levels 5 and 6), with large numbers of students. Eight academics also use e-assessment for postgraduate students, with relatively small numbers (5--25). The final column indicates the percentage of participants who use e-assessment at those levels, with the percentage being the number that uses it at that level over the number of interviewees, namely 68. The percentages total to more than 100, due to the fact that some used e-assessment at more than one level.

Table 3: Levels where e-assessment is adopted -- Interviews

Level of Study	Number of students in cohorts	Adoption rate (number of interviewees who use at that level)	Percentage of interviewees
First year	>250	44	67
Second year	200--500	12	19
Third year	40--120	10	15
Postgraduates	5--25	8	12

The use of different types of multiple-choice questions in electronic assessment

Only 11.7 per cent uses e-assessment for all the assessments in a module, while 29.6 per cent uses it for three to five assessments per cohort, some of them formative. In 45.1 per cent cases of modules investigated, the contribution of e-assessments to final marks, was 11-- 40 per cent, while 11.7 per cent of e-assessments contributed over 80 per cent to final marks.

51.5 per cent of academics did not require any formal training on the tool adopted, but were self-taught. The other 48.5 per cent attended some form of training or orientation. Most departments provide no technical support for academics who adopt e-assessment, and 36.8 per cent of participants solved problems by troubleshooting the issues themselves. University support was given to 29.4 per cent of academics, while in 25 per cent of cases, developers provided support. With regard to administering assessments, 47.1 per cent of the academic interviewees did it themselves, while 52.9 per cent had assistance from administrators or invigilators.

The interviews gleaned qualitative data about interviewees' perceptions of the benefits of e-assessment. The researcher studied the interview transcripts to identify themes and patterns, some of which corresponded with findings from the literature, while others corresponded with those from novel. They fall into seven categories and are supported by anonymised quotations.

### 1. *Feedback to students*

Fast turnaround time is a prime advantage. Detailed constructive feedback can be provided in formative assessments. The correct answer is given, possibly supplemented by information such as page references, hyperlinks to relevant resources, as well as diagrams. Some interviewees cautioned this enrichment feedback. It should be accessible to the student both in the assessment venue and in a portable form, such as a printout or download to a USB. Moreover, the feedback should be equitably provided to all test-takers, including those who got right answers.

### 2. *Consistency*

In automated assessment, there is no subjectivity, bias or impact of a human assessor's emotional or physical state. Nor do different markers assess the same test. In contrast, the responses are judged impartially and objectively by a computer program.

### 3. *Reduced uncertainty*

Students are not left to wonder for days or weeks about their mark, since results are available rapidly. A quotation from Respondent 34 relates to Points 2 and 3:

R34: To get proper turnaround and mark papers/assignments quickly and efficiently and return them to students is very difficult. Also, it is often not fair, because we used to hire a variety of tutors to mark. Online assessment helps to achieve consistency.

### 4. *Assessment frequency*

E-assessment allows for more frequent assessments without additional marking. However, there is a great deal of work and pilot testing when the initial question bank is created. This work continues over successive years, as questions are improved and added. Where enriching feedback is created and included, there is more time-consuming work for the academic.

R28: The main motivation for using it is so I can do continuous assessment on a very regular basis with large numbers of students, without the additional burden of marking.

### 5. *Question sharing and re-use*

Questions banks can be created and questions can be re-used over time. Where more than one higher-education institution offers a course with similar subject matter, questions can be shared.

R8: ... not reinventing the wheel -- sharing of questions and collaboration with peers; question banks.

### 6. *Accessibility*

There can be time- and location independence. Assessments and practice can run on the Web with 24-7 flexibility and in a location of the learner's choice. This is termed 'anytime-anywhere-access'.

R26: Students can attempt the quizzes anywhere because it is linked to the LMS which is accessible through the Internet.

R14: Class tests can be taken from anywhere.

The use of different types of multiple-choice questions in electronic assessment

Where assessment is true formative assessment that does not contribute to marks can be done by the student at any place -- a computer laboratory or kiosk, his/her home, or at his/her workplace. Where it does contribute, caution must be exercised in allowing test-takers to do assessments anytime-anywhere, since there is no assurance it is their own unaided work. One way of countering this is by timed pauses. Questions remain on the screen for a fixed time, which is insufficient to search for the answer in a book. However, it does not prevent the situation where the student has someone on hand to help answer questions. Assessment that contributes towards the mark should be done in official, monitored venues, with students doing it simultaneously or in sessions to manage large numbers of students.

### *7. Student engagement*

This overlaps with accessibility, since it relates to students answering questions for practice or learning gain. One interviewee mentioned 'challenge questions', which stimulate students to engage beforetime with the upcoming course material.

## **CONCLUSIONS**

This study has established that the usage by South African academics of the various types of multiple-choice questions is concentrated on Multiple-Choice: single and multiple response, True/False and Fill-in-the Blank questions. The more novel question types such as Extended-Matching Items, True/False with explanation, Diagram/Video Clips, Simulation, Multiple Choice: single and multiple response were found to be relevant for assessing higher-order thinking skills.

For first-year, adoption is concentrated on MCQ (single response) and True/False, while at second, third and honors levels, Diagram/Video Clips and Simulation are the most used.

Benefits associated with e-assessment include automated feedback, consistency, rapid scoring, reduced uncertainty, assessment frequency, question sharing and re-use, as well as student engagement, all of which can support assessment practices in open distance learning, where large numbers of students are widely dispersed.

Future research aims at establishing a framework of requirements for, and evaluation of, online assessment systems and tools.

## REFERENCES

NB. Authors to please include the Bloom et al (1956, 1964), Passey 2010 in the reference section)

- Alton, R. 2009. *How to write Multiple Choice Questions (MCQs)*. Available at: <http://rlo-cetl.blogspot.com/2009/05/how-to-write-multiple-choice-questions.html> (accessed 14 June 2011).
- Bani-Ahmad, S. and A. Audeh. 2010. Re-engineering of multiple-choice exam-form production tools: Cost-effective and quality-assurance approach. *Paper presented at the 2010 International Conference on Educational and Information Technology (ICEIT), 17--19 September 2010. China: Chongqing.*
- Belton, M. and J. Kleeman. 2001. New developments in perception designed to meet changing face of CAA. Paper presented at the 5<sup>th</sup> CAA Conference, Loughborough: Loughborough University.
- Beullens, J., B. Van Damme, H. Jaspert and P. J. Janssen. 2002. Are extended-matching multiple-choice items appropriate for a final test in medical education? *Medical Teacher* 24(4): 390--395.
- Chiheb, R., R. Faizi and A. E. Afia. 2011. Using objective online testing tools to assess students' learning: Potentials and limitations. *Journal of Theoretical and Applied Information Technology* 24(1): 69--72.
- Costagliola, G. and V. Fuccella. 2009. Online testing, current issues and future trends. *Journal of e-Learning and Knowledge Society (Je-LKS)* 5(3): 79--90.
- De Villiers, M. R. 2005. e-Learning artefacts: Are they based on learning theory? *Alternation* 12(1b): 345--371.
- Fenderson, B. A., I. Damjanov, M. R. Robeson, J. J. Veloski and E. Rubin. 1997. The virtues of extended matching and uncued tests as alternatives to multiple choice questions. *Human pathology* 28(5): 526--532.
- Fielding, A. and E. Birmingham. 2003. Tools for computer aided assessment. *Learning and Teaching in Action* 2(1).
- Kadhi, T. 2004. *Online assessment: A study of the development and implementation of a formative online diagnostic tool in a college developmental mathematics course*. Paper presented at the annual meeting of the Southwest Educational Research Association, 5--7 February 2004. Dallas: Texas A and M University.
- Karl, M., F. Graef, S. Eitner, M. Wichmann, S. Holst and N. Beck. 2011. Student attitudes towards computer-aided testing. *European Journal of Dental Education* 15(2): 69--72.
- Khare, A. and H. Lam. 2008. Assessing student achievement and progress with online examinations: Some pedagogical and technical issues. *International Journal on e-Learning* 7(3): 383--402.
- Khedo, K. K. 2005. *Computer-assisted assessment system at the University of Mauritius*. Paper presented at the IEEE 3rd International Conference on

- Computational Cybernetics (ICCC), 13--16 April 2005. Mauritius: Hotel Le Victoria.
- Luckett, K. and L. Sutherland. 2000. Assessment practices that improve teaching and learning. In *Improving teaching and learning in higher education: A handbook for Southern Africa South Africa*, ed. S. Makoni, 98--130. Johannesburg: Witwatersrand University Press and The Higher Education Research and Development Society of Australasia.
- Maurice, S. A. and R. L. Day. 2004. Online testing technology: Important lessons learned. *International Journal of Engaging Education* 20(1): 152--160.
- McAlpine, M. 2002. *Principles of assessment*. Available at: <http://www.caacentre.ac.uk/dldocs/Bluepaper1.pdf> (accessed 11 October 2011).
- Mitkov, R., L. A. Ha and N. Karamanis. 2006. A computer-aided environment for generating multiple-choice test items. *Natural Language Engineering* 12(2): 177--194.
- Pittenger A. L. and J. L. Lounsbery. 2011. Instructional design and assessment: Student-generated questions to assess learning in an online orientation to pharmacy course. *American Journal of Pharmaceutical Education* 75(5): 1--8.
- Scalise, K. and B. Gifford. 2006. Computer-based assessment in e-learning: A framework for constructing 'Intermediate Constraint' questions and tasks for technology platforms. *Journal of Technology, Learning, and Assessment* 4(6): 1--44.
- Ventouras, E., D. Triantis, P. Tsiakas and C. Stergiopoulos. 2011. Comparison of oral examination and electronic examination using paired multiple-choice questions. *Computers and Education* 56(3): 616--624.
- Walker, D. J., K. Topping and S. Rodrigues. 2008. Student reflections on formative e-assessment: Expectations and perceptions. *Learning, Media and Technology* 33(3): 221--234.