Peer-Reviewed Article

Computer Assisted Assessment (CAA) in Higher Education: Multi-text and Quantitative Courses

Yaron Ghilay, Ph.D.
The NB School of Design and Education, Haifa, Israel

Abstract: The aim of the study was to examine the effectiveness of Computer Assisted Assessment (CAA) regarding multi-text courses compared to quantitative courses in higher education. The research was based on three samples of students ($n_1 = 20,\ n_2 = 29,\ n_3 = 27,\ n_{total} = 76$) who studied three courses, assessed by CAA (multi-text and quantitative). The learners were asked to answer an online questionnaire to evaluate their perceptions towards this method of assessment. Besides, the courses' grades were compared to find out if there are significant differences among them. The findings indicate that CAA has significant advantages for both types of courses regarding its effectiveness, integrity, experience, and flexibility. Furthermore, the grades are stable over time, and are independent on the type of course. Thus, it is recommended to adopt CAA in faculties of higher education teaching both quantitative and multi-text courses.

Keywords: CAA: Computer Assisted Assessment, quantitative course, multi-text course, Moodle assessment tool

Introduction

General Background

As assessment is critical to learning (Brown, Bull & Pendlebury, 1997), considerable pressure has been placed on institutions of higher-education to measure learning outcomes more formally and more often (Ghilay, 2017; Brown et al., 1997; Farrer, 2002; Laurillard, 2002). Computer-assisted assessment (CAA), has the potential both to ease the assessment load and to provide innovative and powerful assessment tools (Brown et al., 1997; Bull & McKenna, 2004). Furthermore, as use of ICT increases, 'inherent difficulties in teaching and learning online but assessing on paper' may arise (Bull, 2001; Bennett, 2002).

Objective test questions (multiple-choice, true/false, or numeric answers) which require the choice of a predetermined answer are the most common format in CAA. However, other types of questions are also possible, such as essay, matching, matrix question, drag and drop into text/image, selection of missing words, etc. (Ghilay, 2017). A major concern on the part of both academics and quality assurance staff (Bull, 1999; Warburton & Conole, 2003) is whether multiple-choice questions can assess higher-order learning outcomes (Pritchett, 1999; Davies, 2002). The most optimistic view states that if sufficient care is taken in the construction of questions (Farthing & McPhee, 1999; Duke-Williams & King, 2001), item-based testing may successfully examine the full range of learning outcomes. This can be done using questions that students cannot answer by relying solely on memory. Good questions should examine understanding and the ability of students to apply, analyze, evaluate, and create.

CAA is not a new approach; it has been rapidly developing over the last decades in schools, universities, and other institutions, offering educational and technical sophistication that includes simulations and multimedia-based questions which are not feasible in paper-based

assessments (Bull & McKenna, 2004). As the numbers of students are increasing while financial resources are decreasing, objective tests may offer a way to cut costs.

CAA systems must be tailored so students receive advanced training before taking part in summative evaluations so as to avoid the danger that CAA evaluates competence in the use of information technology rather than the knowledge of the course content.

Possible advantages of CAA are (Bull & McKenna, 2004):

1. Increasing the frequency of evaluation which may motivate students to practice their skills and learn more
2. Broadening the range of knowledge assessed
3. Increasing feedback to students and lecturers
4. Extending the range of assessment methods
5. Increasing objectivity and consistency
6. Decreasing marking loads
7. Boosting administrative efficiency

A study by Ghilay and Ghilay (2012) reveals that CAA has significant advantages over traditional assessment. On the other hand, Schoen-Phelan & Keegan (2016) have found no statistically significant increase using CAA versus traditional Paper-Based Assessment (PBA). According to Schoen-Phelan & Keegan (2016), students did not appear to have gained a higher level of satisfaction conducting tasks on a computer. One of the reasons given was that most exams at the institution are PBA; consequently, assessments should prepare for the related exams, and should also be paper-based.

Ghilay (2017) claims that CAA is known for its credibility, flexibility, and ability to deal with large amounts of students being tested simultaneously. Elimination of the use of paper eliminates the time and expense of sending questionnaires out for printing and reproduction of scanning pages of answers. With computerized assessment, all these tasks are done via digital means. Due to the growing popularity of a mobile, lightweight computing, it is now possible to conduct computerized tests in regular classrooms (not only in a computer lab). Use of the institutional Wi-Fi with students connected via different private devices - laptops, tablets, and even smartphones - is feasible. A Moodle application allows all registered users to be identified and examined by entering username and password.

The Moodle assessment tool is sophisticated, regularly updated, and includes lots of options. There are various tools for building computerized tests. Therefore, serious preparation toward universal computerized assessment should be made, with instructors learning the details of the application and constructing question banks and exams (Ghilay, 2017).

Characteristics of CAA in Moodle

Computerized assessment via Moodle has unique characteristics requiring precise definition depending on the desired educational objectives (Ghilay, 2017):

1. Time and duration. The date of the test can be restricted (or not) to a specified time window. Additionally, the duration can be set so that the timer starts only from the beginning of the test. Examinees can submit the test on their own initiative before the end of the time limit, but automatic submission occurs as soon as time runs out. The duration of an exam can also be unlimited, an option that is appropriate for formative assessment or practice.

2. The number of answering attempts. The system allows lecturers to determine the number of response attempts allotted to examinees. If this is a summative evaluation (a final

exam), only one attempt will be permitted. However, if it is a practice exam, examinees may be allowed to submit answers several times. Another option is to allow unlimited attempts – with respondents being able to go back and restart the test/exercise as they wish. With multiple response attempts, the system can allow new attempts to be built upon preceding answers (or not). This option is suitable for practicing. If students prefer every new attempt to start from scratch, all answers are blank. Instructors may give students the option to choose the response method they prefer.

If questions are randomly selected from a databank, it is most likely that examinees will have different items than on previous exams or that questions appear in a different order. Moreover, because in Moodle, duplicating tests is very easy, a practice exercise can be made in two versions: in one alternative, new attempts are built upon preceding answers and in the other one, every new attempt starts from scratch.

Another relevant consideration in multiple response options is determining how the final score should be calculated: highest grade, average grade, or last attempt. Also, a minimum wait period between two successive attempts can be set so as to encourage students to study the material before answering another exercise or test.

3. Place limitations. Since Moodle is an internet platform, any authorized user can respond to the test/exercise on any computer connected to the net. However, there may be occasions (particularly those of summative assessment) where it is necessary to limit the place from which students are allowed to respond. For example, students must take the exam in specific rooms in the institution and not in their homes. Therefore, specific IP addresses from which students are allowed to register to take the test can be set. Anyone logging in from an

address that does not belong in the institution network (including cellular networks) will be automatically blocked.

4. Producing different questionnaires. The computerized system has unlimited capacity to produce a large collection of profiles for a given test at a given time enabling each and every questionnaire to be different. This feature is key to maintaining exam integrity. With manual methods, it is hard if not impossible to approach such differentiation. Configuration of the question bank into topics and subtopics enables random selection of X questions out of Y items in the database. With big databases, more sophisticated test questionnaires can be created. For example, separate databases can be set up for voluntary exercises and for compulsory exercises (tests). In such a reservoir, random selection of questions by topic ensures no overlap among issues relating to different tasks (exams, practice exercises, and so on).

Another option in creating differentiation in test questionnaires is control of question order and in multiple-choice questions, the order of options. While the order can be predetermined, random order is preferred.

5. Setting different parameters for different examinees or groups. The system allows significant degrees of freedom for meaningful exceptions to the general test conditions. Deviation from the general settings can be adding an additional password to certain students for entering the test, changing the start/stop time, the time limit, or the number of attempts allowed. For example, a group of students with learning disabilities for whom the time limit could be three hours instead of two. In other words, in the same test, there may be different deadlines for different students. Such a differentiation can also exist on an individual basis - a particular student can have various restrictions designed just for him/her. For instance, it is possible to give

a particular student a time extension for submission of an exercise due to special personal circumstances, a relief which does not pertain to the rest of the class.

6. Grade management. Even in this area, computerized systems have great flexibility and save on manual work. Assessment of "objective test questions" is performed automatically and grades are retained within the system. Lecturers may update any grade they wish. Also, any item can have a weight (the default is 1) for automatic weighting of the final score. In the case of open-ended questions where manual evaluation is required, scores can be automatically weighted for final score calculation.

Lecturers can determine whether and when test scores will be posted as well as different types of feedback received. Also, the class grade sheet including a collection of exercises or tests (all or part of them) can be exported to common data management formats such as Excel which might be helpful for further processing, reporting, or external storage.

Instructors have the option to review the results of the test of every student, read all the answers, seeing what was done right and what was done wrong. So, without any paperwork, full and detailed documentation of all student responses is possible. Based on these digital reports alone, it is possible to have a discussion with students concerning test results.

7. Restricting access to the test/exercise. Lecturers can set advanced restrictions on access to tests/exercises, for example, access to late-testing questionnaires for those who have not yet been tested at all. This may be relevant in situations where some students are retaking the test while others are just making their first attempt. Without any human intervention, it is possible to ensure that each group of students is tested only on a suitable questionnaire, enabling various tests to be held at the same time.

8. The variety of questions. It is believed that computerized assessment deals only with multiple-choice questions, but actually this is not true. Moodle enables more than 30 different types of questions, the most common of which are:

Multiple-choice: An unlimited number of distractors can be mixed and numbered in different forms (or not numbered at all). Lecturers can decide whether students must select only one answer or are allowed multiple answers. If there is only one right answer, wrong answers can get a zero or a partial score. When required to select multiple answers, each correct answer has to receive a partial mark so that the total grade adds up to 100%. Other distractors are scored as zero.

Numerical: In numerical questions a final numeric answer is required with examinees writing the correct answer in a box at the bottom of the question. It is also possible to require that students include measurement units for the numerical values. Permissible deviation from the exact value can also be defined.

Matching: A matching question displays two lists of items on either side of the screen. Students select an answer from the list of items on the right side that match an item listed on the left side.

Entering missing words in a paragraph: Examinees are asked to add missing words in a text by choosing from a menu with several options.

General description of the study

The aim of this three-year study was to examine the effectiveness of Computer Assisted Assessment (CAA) of multi-text courses compared to quantitative courses and to explore whether there are differences between the CAA of both types of courses.

Students who studied the following three courses were examined:

1. Entrepreneurship (distance learning, fourth year students).
2. Fundamentals of Technology Management (distance learning, second year).
3. Mathematics for Business Administration (face-to-face learning, first year).

All three courses included computerized Moodle exercises and a computerized final exam. All the students who participated in the study, studied in the Department of Management and Economics at the NB School of Design and Education, Haifa, Israel.

The computerized final exams had the following characteristics:

Entrepreneurship (multi-text): the test covered 9 topics and included 30 multiple-choice textual questions. The 30 questions were randomly selected for each topic separately, out of a question bank of 45 questions.

Fundamentals of Technology Management (multi-text): the test covered 8 topics and included 30 multiple-choice textual questions. The 30 questions were randomly selected for each topic separately, out of a question bank of 52 questions.

Both exams of multi-text courses (Entrepreneurship and Fundamentals of Technology Management) have the following common characteristics:

1. Each question includes one correct answer and three distractors.
2. Every answer (correct or incorrect) is not too short (about 30-40 words each).
3. The method of generating the answers is based on the following procedure:
    - Wording the right answer (answer no. 1).
    - Copying and pasting the right answer three more times (answers no. 2-4).
    - Converting answers 2-4 to distractors by making a small change (a word or two) that causes the answer to be incorrect.
    - The distractors were tested to be immune to response attempts based on pure

logic rather than knowledge of the course material. The great length of the correct answer and the distractors, as well as the slight differences between them, are intended to make it difficult for the examinees to reach the correct answer by improper means.

- The fact that the correct answer is always the first does not have any effect because the order of answers (for each question) is set to randomize for each response attempt.

4. The questions were presented to the examinees in the order of topics on separate pages (topic 1: page 1, topic 2: page 2, and so on). Each student received random selected questions for each topic.

Mathematics for Business Administration (quantitative): The quantitative exam has the following characteristics:

1. The test includes four topics divided into twelve subtopics that are covered by twelve quantitative questions.
2. The twelve questions are randomly selected for each subtopic separately, out of a question bank of 60 questions. This means that each question is randomly selected from 5 options.
3. There are four types of questions as follows:
    - Multiple-choice: Each question includes one correct answer and four distractors.
    - Matching questions.
    - Drag-and-drop matching questions: This type of question (as well as the previous) includes some sub-questions. Students can get a partial score for a

particular question, if not all sub-questions are answered correctly.

- Numerical questions (examinees are required to enter a numeric answer).

Dealing with such types of quantitative questions may be problematic in terms of the integrity of the exam. Unlike a multi-text course, the short length of the answers in a quantitative test may make it easier for the examinees to reach the correct answer by improper means. To overcome this obstacle, it was necessary to create at least 5 different questions for each subtopic, so the probability that two students sitting nearby would receive the same single question would be as low as 0.2 (one to five). As for the complete exam, the total number of options to create different test questionnaires is huge ($5^{12} = 244,140,625$).

Methodology

The study examined students' attitudes towards CAA in three courses, divided into two main categories: multi-text and quantitative. The same lecturer prepared all the exams and conducted the three courses.

The research questions

The research questions aimed at measuring the effectiveness of CAA in higher education and finding differences (if any) between multi-text and quantitative courses.

The following research questions were phrased:

1. Based on students' views, what are the advantages/disadvantages of CAA regarding multi-text courses versus quantitative courses?
2. Are there significant differences between the means, dispersion, and distribution of test grades in multi-text courses compared to quantitative courses?

Population and Samples

Population. The population addressed through the study included all learners in Israeli institutes of higher education who are examined by CAA.

Samples. Three samples that have been examined are presented in Table 1.

Table 1

The study samples

No.	Course	Year	Category	Way of learning	Sample size	Rate of response
1	Entrepreneurship	2015-16	Multi-text	Distance	20	100% (20/20)
2	Fundamentals of technology management	2016-17	Multi-text	Distance	29	93.55% (29/31)
3	Mathematics for Business Administration	2017-18	Quantitative	Face-to-face	27	93.10% (27/29)
	Total				76	

Tools

First research question. Respondents were asked to answer an online five-point Likert scale questionnaire (1-strongly disagree, 2-mostly disagree, 3-moderately agree, 4-mostly agree, 5-strongly agree) consisting of 24 items and an open ended question. The questionnaire examined students' attitudes toward the first research question and it was anonymous.

Second research question. Three courses' mean grades and variances were compared to find out if there are significant differences among them. Moreover, a comparison of grades' distribution presented in the following 7 grade categories was undertaken: 0-40, 41-50, 51-60, 61-70, 71-80, 81-90, and 91-100.

Data Analysis

First research question. The following four factors were examined: Test effectiveness, test integrity, test experience and test flexibility. Table 2 summarizes the four factors, the items composing them and the reliability. For each factor, a mean score was calculated (including standard deviation). One Way ANOVA was conducted to examine significant differences among the above three courses. Paired Samples T-test was undertaken as well to check significant differences between pairs of factors ($\alpha \leq 0.05$).

Second research question. One Way ANOVA and Levene Test of Homogeneity of Variances were conducted for checking if there are significant differences among the three courses' mean grades and variances. Chi-Square Test was undertaken as well for checking if grade categories distribution depends on the type of course ($\alpha \leq 0.05$).

Table 2

Factors and reliability.

Factors	Questionnaire's Questions
Test effectiveness (Alpha=0.788)	The test measures the level of my knowledge accurately: The test covers the required course material well. The test assesses basic learning objectives (knowledge and understanding). The test assesses high learning objectives (implementation, analysis, etc.). The test covers broad areas of the course. The test is objective and consistent.
Test integrity (Alpha=0.757)	Getting help from other examinees is difficult. There is no chance of exam questions being leaked. Examinees receive different test questionnaires. Test integrity is carefully maintained.
Test experience (Alpha=0.758	I enjoyed the exam. I felt comfortable during the exam. Receiving my score at the end of the exam was an advantage. I am sure my answers reached the lecturer in a proper manner. It is convenient for me to give answers on a computer screen. It is a convenient method for updating answers that I want to change before submission. I'm not worried about the exam. It is easy to concentrate on questions displayed on a computer screen. The test includes a variety of assessment methods. The time limit does not disturb my concentration. I can appeal the examination results.
Test flexibility (Alpha=0.721)	I can get multiple opportunities to be tested. There are many opportunities to improve my grade. The lecturer can be flexible concerning the dates of exams.

Results

First research question

Table 3 presents the mean scores of three samples:

Table 3

Samples' mean scores

Factor	Sample	N	Mean	S.D
Test effectiveness	Entrepreneurship	20	4.49	.40
	Fundamentals of technology management	29	4.55	.52
	Mathematics for Business Administration	27	4.46	.52
Test experience	Entrepreneurship	20	4.21	.49
	Fundamentals of technology management	29	4.18	.43
	Mathematics for Business Administration	27	4.42	.44
Test flexibility	Entrepreneurship	20	4.20	.63
	Fundamentals of technology management	29	4.38	.55
	Mathematics for Business Administration	27	4.02	.68
Test integrity	Entrepreneurship	20	4.29	.53
	Fundamentals of technology management	29	4.19	.73
	Mathematics for Business Administration	27	4.56	.55

One-way ANOVA was conducted to determine whether there were significant differences among the mean scores of all samples for the above factors. The following results were obtained:

1. Test effectiveness: $F_{(2,73)} = .274,\ p = .761$

2. Test experience: $F_{(2,73)} = 2.163,\ p = .122$

3. Test flexibility: $F_{(2,73)} = 2.311,\ p = .106$

4. Test integrity: $F_{(2,73)} = 2.680,\ p = .075$

The above findings indicate that no significant differences were found among the means of all the samples, for all factors (ANOVA, $\alpha \leq 0.05$). Thus, the mean factors for all these three samples together are shown in Table 4.

Table 4

Mean factor: Three samples together (descending order)

Factor	N	Mean	S.D
Test effectiveness	76	4.50	.49
Test integrity	76	4.35	.64
Test experience	76	4.27	.46
Test flexibility	76	4.21	.63

According to Table 4, the effectiveness of the test was rated with a very high score (4.50). The significance of this finding is that the examinees claim that a computerized test accurately measures their knowledge, it covers the course material thoroughly, it evaluates basic learning goals (knowledge and understanding) as well as high learning objectives (implementation, analysis, etc.). Moreover, according to students' views, this type of test covers broad areas of the course and is objective and consistent. The other three factors are also highly rated:

Test integrity (4.35): Getting help from other examinees is difficult, there is no chance of exam questions being leaked, examinees receive different test questionnaires, and thus, the test integrity is carefully maintained.

Test experience (4.27): The students enjoy the experience, feel comfortable, are not concerned, concentrate, and claim that a variety of evaluation methods is an advantage.

Test flexibility (4.21): The examinees can have many opportunities to take the exam and improve their score, and there is considerable flexibility with regard to examination dates.

Based on Paired samples T-Test, there were significant differences between the following pairs of factors:

Test effectiveness (4.50) and Test experience (4.27): $t_{(75)} = 4.134, p = 0.000$

Test effectiveness (4.50) and Test flexibility (4.21): $t_{(75)} = 3.590, p = 0.001$

Between all the other pairs of factors, there were no significant differences. The meaning of these findings is that the effectiveness of the computerized examination is considered significantly higher than its experience and flexibility.

The open-ended question reinforces the closed items and gives them more validity as presented in the following quotations of respondents:

Entrepreneurship:

"It was a great way to be examined. I think that computerized exams have only advantages."

Fundamentals of Technology Management:

"Very good method. In my opinion, it can completely replace the traditional paper exam."

"This method is very effective and convenient."

Mathematics for business administration:

"I prefer computerized tests because they are more convenient and easier to read."

"It was convenient for me and easier to use the computer because we had already learned the material on the computer and we did a lot of computerized exercises so we got used to it."

"A computerized test is, in my opinion, a great advantage because it is convenient, efficient, flexible and provides an excellent testing experience. I prefer this form of test over traditional paper and pencil tests."

Second research question

Based on One Way ANOVA ($\alpha \leq 0.05$) and Levene Test of Homogeneity of Variances, no significant differences were found between the mean grades of the three courses ($F_{(2,77)} = 1.326,\ p = .271$) and their variances ($Levene\ F_{(2,77)} = 2.049,\ p = .136$).

The mean grades and S.D for the three courses are shown in Table 5.

Table 5

Mean grades

Course	N	Mean	S.D
Entrepreneurship	20	84.03	8.83
Fundamentals of technology management	31	87.74	8.58
Mathematics for Business Administration	29	84.29	10.87
Total	80	85.56	7.57

Table 6 presents the distribution of grades divided into 5 categories (the first two categories, 0-40 and 41-50 are empty).

Table 6

Distribution of grades

Course		Grades' categories					Total
		51-60	61-70	71-80	81-90	91-100	
Entrepreneurship	Frequency	0	1	6	9	4	20
	%	0.0%	5.0%	30.0%	45.0%	20.0%	100.0%
Technology	Frequency	0	2	5	14	10	31
	%	0.0%	6.5%	16.1%	45.2%	32.3%	100.0%
Math	Frequency	1	1	7	7	13	29
	%	3.4%	3.4%	24.1%	24.1%	44.8%	100.0%
Total	Frequency	1	4	18	30	27	80
	%	1.2%	5.0%	22.5%	37.5%	33.8%	100.0%

According to Chi-Square test, it was found that the distribution of grades does not depend on the type of course ($\chi^2_{(8)} = 7.484,\ p = .485$).

Discussion

Examinations in institutions of higher education are an important task in any academic course. Running exams while maintaining high effectiveness, integrity and providing students with appropriate experience and flexibility is a major challenge for the academic staff. Technology may be very useful and efficient for the preparation, administration and evaluation of exams (Ghilay, 2017). From the perspective of institutional management, it is quite clear that CAA has great advantages. It is paperless and can be very helpful for easily maintaining the time frame. Moreover, such technology is useful for producing countless versions of questionnaires, automatically, instantly and efficiently. CAA has also an advantage in test checking and scoring (Ghilay, 2017). However, it was unclear whether CAA is not only efficient and convenient for administration needs, but is also effective, maintains integrity and provides students with successful testing experience and flexibility, for different types of courses.

The findings show that there are no significant differences between all the courses examined. This means that regardless of the type of course, according to learners' views, CAA

has significant advantages both in multi-text as well as in quantitative courses: the effectiveness of the assessment process is outstanding (4.50), test integrity is highly preserved (4.35), examinees are having a very good experience (4.27), and the complete process is considerably flexible (4.21). According to students' views, it can be concluded that running tests as done in this study is a great advantage.

Besides, additional findings of real outcomes, show that during three years, grades' means, dispersion and distribution are stable over time, and are independent on the type of course or cohort. These findings indicate that CAA has a stable ability to make a proper distinction between students' achievements.

Conclusion

Due to the successful contribution of CAA to the process of student evaluation in multi-text and quantitative courses, it is recommended to adopt this method in higher education. To do this in practice, it is necessary to create and deliver training programs so that the academic staff will be familiar with the principles of CAA, methods, and pedagogy. These principles, as well as additional skills required for online learning, can be purchased on the basis of the TMOC (Training for the Management of Online Courses) model (Ghilay, 2017; Ghilay & Ghilay, 2014). It should be stressed that practical implementation of CAA in higher education is based primarily on the personal knowledge of faculty rather than investing in infrastructure or expensive development projects.

References

Bennett, R. E. (2002). Inexorable and inevitable: the continuing story of technology and assessment. *Journal of Technology, Learning and Assessment 1*(1).

Brown, G., Bull, J. & Pendlebury, M. (1997). *Assessing student learning in higher education* (Routledge, London).

Bull, J. (1999). Update on the National TLTP3 Project The Implementation and Evaluation of Computer-assisted Assessment, in: M. Danson (Ed.), *3rd International CAA Conference,* Loughborough University, 16–17 June 1999.

Bull, J. (2001). TLTP85 implementation and evaluation of computer-assisted assessment: final report.

Bull, J. & McKenna, C. (2004). *Blueprint for Computer-Assisted Assessment*. RoutledgeFalmer, NY.

Ghilay, Y. (2017). ODL: Online distance learning of quantitative courses in higher education. *Advances in Social Sciences Research Journal, 4*(18), 62-72. https://doi.org/10.14738/assrj.418.3698

Ghilay, Y. & Ghilay, R (2014). TMOC: A model for lecturers' training to management of online courses in higher-education. *Journal of Educational Technology, 11*(2), 6-16.

Ghilay, Y. & Ghilay, R. (2012). Student evaluation in higher education: A comparison between computer assisted assessment and traditional evaluation. *Journal of Educational Technology, 9*(2), 8-16.

Davies, P. (2002). There's no confidence in multiple-choice testing, in: M. Danson (Ed.) *6th International CAA Conference,* Loughborough University, 4–5 July 2002.

Duke-Williams, E. & King, T. (2001). Using computer-aided assessment to test higher level learning outcomes, in: M. Danson & C. Eabry (Eds.) *5th International CAA Conference,* Loughborough University, 2–3 July 2001.

Farrer, S. (2002). *End short contract outrage*. MPs insist, Times Higher Education Supplement.

Farthing, D. & McPhee, D. (1999). Multiple choice for honours-level students?, in: M. Danson (Ed.) *3rd International CAA Conference,* Loughborough University, 16–17 June 1999.

Laurillard, D. (2002,). Rethinking university teaching a conversational framework for the effective use of learning technologies (2nd ed.). RoutledgeFalmer, London.

Pritchett, N. (1999). Effective question design, in: S. Brown, P. Race & J. Bull (Eds.) *Computer assisted assessment in higher education.* London, Kogan Page.

Schoen-Phelan, B. and Keegan, B. (2016) Case Study on Performance and Acceptance of Computer-Aided Assessment. *International Journal for e-Learning Security*, 6(1).

Warburton, W. & Conole, G. (2003). CAA in UK HEIs: The state of the art, in: J. Christie (Ed.) *7th International CAA Conference,* University of Loughborough, 8–9 July 2003.