

Computer-Based Assessment: Sounds Easy, is it Really?

Margarida Amaral¹, Hugo Ribeiro²

^{1, 2} Universidade do Porto, Porto, Portugal

¹ mamaral@reit.up.pt

² hribeiro@reit.up.pt

Abstract: There is an ongoing project on the University of Porto (U.PORTO) that aims the association of two didactical tools of information and communication technologies: e-learning and computer-based assessment (CBA).

Implementing CBA at an institutional level raises several issues. Supporting staff, if existing, have to address and solve some hard challenges: pedagogical training in new assessment methods, technical problems in the implementation process and organizational problems. These last are particularly complex in a university campus like U.PORTO: large and spread. The e-Assessment topic is of general interest for those who work in e-learning, specially now in the era of Web 2.0, where many questions arise due to so many different approaches of methodology, strategies and evaluation. U.PORTO is no exception. The need of being more accurate in the assessment process when using the web, even if the teachers use a blended-learning approach, is real and work is being developed in order to try to respond to this need.

Although the e-learning Unit (GATIUP) works towards different assessment methods, this particular paper focus on the implementation of online tests on campus using the University LMS – Moodle.

We will explore the experience acquired in implementing large scale high stake examinations in different faculties: Pharmacy, Economics, Medicine, Dentistry, Humanities. On the school year of 2008/2009, 1548 students did their high stake exams online; in 2009/2010 only in the first semester 869 students were evaluated and graded on Moodle. Also, test analysis has been conducted with interesting information about the items and the quality of the overall exam.

It will also be presented the strategies that U.PORTO is engaging to overcome some of the difficulties found in the implementation of such project.

Keywords: assessment, multiple-choice questions, classical test theory, moodle

1. Introduction and Context

The main purpose of this paper is to give a brief description of an ongoing project at the University of Porto, which aims to articulate two areas of education and technology: e-learning and computer-based tests. The current situation of the University concerning the global e-learning Project is described below. The focus will mainly be on the outcomes and not particularly in the implementation strategy as it is not pertinent for the understanding of our current project on computer-based assessment - CBA. The main focus however is the identification of some constraints that arise from an institutional approach of online assessment implementation.

1.1 U.PORTO e-Learning Project

1.1.1 Implementation and Outputs

Implementing an e-learning project in a large, spread and traditional university like the University of Porto, is a complex task that must involve a considerable number of actors, such as faculty managers, teachers, information systems administrators, librarians, data processing departments and, of course, students. Although we can now easily find reports of experiences carried out in scattered and decentralized large universities like the University of Sydney (Ellis et al., 2007) or in the Manchester Metropolitan University (Morten, Keegan, 2006), it was necessary several years ago for the University of Porto to find its own strategic path to address its specificities in the development of e-learning activities (U.PORTO, 2007), (Soeiro et al., 2000).

In the school year 2003-2004, the U.PORTO began a pilot-project to develop pedagogical contents for a Learning Management System (LMS) that was increasingly integrated with the University Information System - SiGARRA. This project, hereafter "e-LearningUP", has the following general objectives:

- Improve the students' learning conditions outside the classroom;
- Improve the students' autonomy and research capacity;
- Work with the students to obtain from them a more energetic contribution for the learning process;
- Give students the opportunity of having some support documents in advance such that their participation in the classroom becomes more active and consistent;
- Privilege strategies that use student activities' elements, like reading and performing small tasks, which contribute to the training of hypothetical-deductive thought (Falcão, 2006), (Lawson, 2006).
- Quickly publish learning objectives, syllabus, bibliography and other documents of interest;
- Increase the interactivity between the students and the faculty;
- Promote a democratic access to information;
- Increase the opportunity of access to the course contents to dislocated students;
- Acquire experience in techniques of distance education through the Internet.

E-learningUP has particular characteristics, not only because both teachers and students are oriented towards a common final objective in a wide range of areas, but also for the innovative purpose of integrating resources that are already available in an information system with a learning management system. Another pertinent fact is the policy of directing this project towards professors with technological competences as software users. One of the main objectives of this project was the future expansion of contents production to the entire professors' community of the university. This was achieved since the pilot team of teachers involved, constitute a representative sample of the academic community with low technological skills. Periodically during the project, existing results were used to motivate other professors towards the production of contents following proved methodologies, and not to amaze them with overdeveloped multimedia resources, which might become unrealistic and even frightening for those who do not have advance technological knowledge.

We would like to point out the fact of the teachers' e-content development and instructional design on the LMS was, and still is, guided but not modeled by GATIUP.

Three distinct platforms were already used – LUVIT, WebCT (both Campus Edition and VISTA) and Moodle. The LUVIT platform was abandoned in 2006, and was replaced by the Moodle platform, which was already widely used by several professors who had engaged in individual initiatives. Moodle became the unique platform of U.PORTO in 2009. The LMS is now a common resource in the university campus which students, technicians and teachers are familiar with.

At the present time all Faculties are involved in the eLearningUP project, which comprises more than 774 on-line courses on Moodle. Some Faculties, like Sciences and Engineering, have their own support offices, – Moodle@FCUP offers 700 on-line courses to students; Moodle@FEUP offers about 1000 on-line courses. The Medical School also runs its Moodle installation mainly used for lifelong learning training courses and post-graduate courses. (U.PORTO website, 2008)

By online course it should be intended a blended-learning support for presential courses, given that the University of Porto is a traditional campus-based institution.

1.1.2 E-learningUP Outcomes

As a product of this project, several pedagogic materials and strategies for their use arose. These products are analyzed by the professors involved and appraised by the entire academic community, allowing for the sustainability of future decisions concerning actions for the continuous improvement of the quality of the universities' courses. Six annual seminars were held, from which books of proceedings, CDs and DVDs containing the presented work were developed. Also, in 2005, U.PORTO created an internal Excellence Award for e-learning. This award intends to distinguish those that use the LMS with a proper pedagogical strategy in their teaching activities during the school year (Senado da Universidade do Porto, 2005). The e-contents developed for support of the course as well as communication tools or other innovative usage of e-learning is taken under consideration.

From the beginning of this process, the GATIUP's philosophy included the teachers' involvement on national and international conferences, national funding projects, as well as trying to include them in internal discussion groups about pedagogic and paradigmatic questions regarding new technologies in education. GATIUP provides several training courses, both online and face-to-face, in different areas that relate technology and pedagogy: using Moodle, online tutoring, strategies for the use of web 2.0 in learning, how to prepare an exam, e-books, e-portfolios, and animations storyboards and many others. With time, a considerable number of teachers began to feel the increasing need to explore a diversity of traits, while maintaining the belief that they would complement and even improve their own experiences in e-learning with students.

This is an interesting evolution of top-down process with a strong component of interest towards a middle-out process.

2. Motivation for Computer-Based Assessment

Keeping in mind what was prior said about the use of electronic tools in teaching and learning, we will introduce at this point of the paper another relevant experience held in the University.

For many decades multiple choice questions (MCQ) or other types of questions that allow automatic grading, were used in exams through the University. Some faculties even have a systematic digitalization by optic reading of all evaluation tests. The transformation of the optic readings into matrixes of answers allows the application of specific software which, through an iterative calculation procedure, adjusts the data to logistic equations of the IRT (Item Response Theory). The data analysis allows the measurement of difficulty and discrimination parameters for each item, as well as amount of information each item complies. Consequently, a continuously growing set of calibrated items has been built. Moreover, the developed work as been used to guide the current teaching activity of the courses. Some of the teachers use Classic Test Theory analysis instead of IRT, other teachers use specific test software analysis and others only rely on Moodle item analysis, these last are the pioneers in CBA at the University.

CBA was a smooth and natural evolution to these teachers: they had accumulated experience given by paper-pencil exams, and they have now the skills and knowledge to work with new technologies in the pedagogical process. If a LMS was already being used for pedagogic activities such as delivery of documents, assignments, discussions why not extend that to summative evaluation?

Many teachers were using quizzes as formative evaluation tool but didn't take the leap into the final exam.

GATIUP had the role of supporting that leap: making it simple.

3. Individual Course Experience

3.1 Implementing CBA in Physiology

After a successful experience, held in 2007-2008 in the Faculty of Pharmacy, that involved more than 300 students doing an online summative exam, other courses decided to embrace this service.

To the purpose of this paper we will focus on this Faculty – Pharmacy, and in a particular course: Physiology. Physiology, is a course from the second year of the Master of Pharmaceutical Sciences. This year the number of students reached 229.

This is a regular course collaborating with GATIUP: b-learning working on Moodle as LMS allowing students to consult all the documents made available by the teacher. The usage of Moodle was rather interactive having activities such as discussion board for peer discussion, assignments and multimedia objects.

It is also important to emphasize that this course ran multiple choices quizzes before this ability appeared on Moodle. The global evaluation was and is held by having a final exam, plus a distributed evaluation percentage. This year the weight of exam was 70% (14 out of 20).

The exam was assembled using Respondus and imported to Moodle in a total amount of 60 questions each one with five options. Negative scoring was not used. With such a great amount of students and a low ratio of PCs/students the only solution was to have several shifts, 76 minutes each. Having 37 working PCs distributed by three rooms, shifts were published alphabetically in a total of 6.

With the local support of a member of GATIUP and the local network manager no security failures occurred.

By the end of the day – almost eight hours later – there were 196 submissions. The final exam grades were shown to the students in the next morning – teacher’s decision despite automatic scoring of Moodle. Our office supports this procedure in summative examinations. Sometimes errors are detected only after the exam delivery and there is a need of regarding certain questions.

The main advantage of this process relies on three main facts:

- No time spent revising answers
- Absence of human mistakes on correction (depending on the exam assembling)
- Possibility of immediate feedback and grading.

We could also point out that it seems harder to cheat, especially because of the room U configuration and the teacher’s surveillance. No need to say that internet access was disabled. On the Moodle side, the system only allows proceeding for those who are accessing Internet through that room’s LAN and the exam has also a password which is never transmitted to the students - GATIUP’s staff goes out to each student and types the password himself.

3.1.1 Results observed

Even though we got some student’s opinions about a greater level of stress on CBA there’s nothing that makes us go back to the paper. Even ecologically this is a best practice. By now a questionnaire regarding online assessment with Moodle is being prepared to study in more detail the students’ evaluation of such tool.

As mentioned before, due to large number of students, the staff and teachers have to organize the room’s occupation in shifts. Does this have impact on students’ grades? Meaning, in which way does the grade varies through the day. And also how have the mean grade evolve since the last exam on paper. We are aware of the great amount of variables involving these numbers and that’s why we intend to continue to trace them through time. Nevertheless, we can have a picture of what’s happening in each exam.

Further research and analysis is being held by the supporting office for New Technologies in Education. We are analyzing several data from different exams in different faculties comparing shifts performance thru the examination day. One course didn’t allow taking any conclusions but now a systematic analysis of these situations is being done.

Teachers are collaborating in an extended analysis that relates paper-pencil tests with computer-based, providing information about previous examinations paper delivered. In Figure1 you find representation of mean of grades in a summative evaluation in the last 3 years for Physiology Course. There are no statistical significant differences between the 2007-08 year – paper exam – and subsequent years – computer-based assessment.

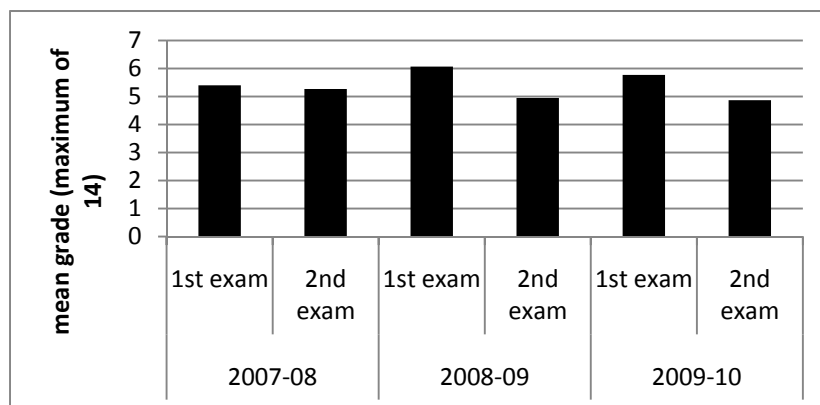


Figure 1: Graph showing the performance of students in different years. 2007-2008 was still paper-pencil and after that exams were online. (Physiology Course – Faculty of Pharmacy of U.PORTO)

There is an intrinsic idea of both students and teachers that the medium influences the final score of the students on an exam. In this particular case we cannot say it has since one can't find a significant difference between grades.

Another problem we have is the question bank: is it diverse and robust enough? Does item exposure occur? Again there were no evidences of such, by analyzing the results of 1st exam (took a few months or weeks before the 2nd exam) and 2nd exam.

4. Institucional Support for CBA Implementation

The trust developed between the supporting unit for New Technologies in Education and the academic community is vital for the implementation of such project. Final assessment is a serious phase in the pedagogic process along a course and it has to be secure and reliable.

Although the supporting team had already a sustained knowledge of educational technology, a hard work on researching and self-training in the area of computer-based assessment (CBA) and computer adaptive testing is being made.

The process to extend from e-learning to CBA looks rather as a natural evolution and from the conceptual point of view doesn't seem to bring special difficulties. However the extension to the entire campus raises scalability issues difficult to solve, of which two are especially concerning:

How to make this process transparent to faculty;

How to assure some specific competencies such as psychometrician expertise and human and physical resources are proper and reliable.

It is not easy to find in the literature a report of a strategy of implementation of a CBA system to a large community of academics. It is common to find detailed descriptions about CBA construction, delivery and analysis with all the mathematical models explicitly shown and explained within the context.

The University of Porto intends to create a structure that contains in itself the possibility for the academics to use CBA in a transparent, secure and reliable way. By this, the work will focus in creating conditions of availability of these tools for every teacher who wants to use the computer as a mean to produce, deliver and analyze a test. It is important that all the models and mathematical background are assured by a psychometrician, who must support the teachers who should not need to run through the theories and principles of item analysis.

The support office for new technologies in education traced several steps in order to pursue this goal: Training Faculty Support Staff - One of the functions of a faculty support office that is often overlooked, even though it is expensive and time-consuming, is the training of its own staff. Faculty support personnel must learn (and buy or find open source solutions) new versions of software in a timely manner, and must frequently add adding new knowledge to their repertoire of skills and understanding.

Training the academic community – building multiple choice questions, assembling a test based on the computer, IRT, Classic Test Theory and other subjects relevant to the understanding and execution of this project. Training of the academic staff in the development of multiple choice questions is a demand of the teachers that work in this project. GATIUP already ministered a 14 hours training course in the development of MCQ. The second edition of this course took place in January 2010 with the presence of an invited expert in Psychometrics that addressed 61 teachers on a workshop. The event was held under the punch line: "Is intuition enough to build an exam?". It was a successful workshop which proved to be an indicator that this project, thru training is already increasing the awareness around questions construction and their quality analysis a priori and a posteriori.

The teachers involved in online testing are now on the calibration phase of their questions. In addition several other teachers that still use paper-pencil in their exams are contacting GATIUP in order to analyse through Classic Test Theory their exams results in previous years. Good indicator that the item analysis is an important tool for a better understanding of an exam quality in the evaluation process.

Analyzing software available on the market - the IMS Global Consortium, an industry and academic consortium, produced the IMS QTI (IMS Question & Test Interoperability Specification) to enable the exchange of questions and test data and also results reports (IMS Global Learning Consortium, 2006). The QTI specification comprises several documents which are being taken under consideration in the decision of choosing assessment software with interoperability of question design. Ferl Becta (Ferl Becta, 2006), Suleman (Suleman, 2003) lists 21 software tools for e-assessment. From that list, only 4 products provide IMS QTI and SCORM support to some extent.

In relation with learning platforms, the scenario of interoperability doesn't get any better. Moodle, which is used at U.PORTO, allows to export IMS QTI 2.0 format but not to import. (Moodle Statistics, 2006), (Moodle Release Notes, 2006) It is rather satisfactory and it allows the perfect articulation between course contents and communication tools with evaluation tools. Students need a single authentication procedure in Moodle to start their exam: there's no need of contacting with other type of software.

We are carefully analyzing offline alternatives that can be synchronized with Moodle. This is a very important aspect because of the Moodle server load during online exams. The number of teachers and students is rising every exam season. Some alternatives were tested with no satisfaction of performance. GATIUP will probably straighten Moodle's architecture in order to allow more concurrent users simultaneously.

Testing Rooms – to perform CBA it is necessary to guarantee rooms with a considerable number of computers prepared for assessment. These rooms with specific characteristics (Pearson Vue, 2007) to provide a distraction-free, secure testing environment with continuous candidate surveillance can be already found in some of our faculties.

With consecutive experiences we expect to identify outcomes that allow us to make adjustments to our strategy in order to succeed with the development of resources for computer-based testing in campus-wide IT systems.

5. Conclusions

CBA is an obvious extension of e-learning networks in academic settings. However, these tools require complex analysis and expensive equipment. The scientific background of CBA, with MCQ or not, is very robust. Validity and reliability issues are also well characterized in many studies.

The large number of students, time spent on the process (since development to classification), lack of adjusted tests, inadequate methods relating to new paradigms of teaching and learning are current problems that U.PORTO academics face. These lead to the engagement in the use of innovative assessment methods.

Online assessment with automatic feedback presents itself as a possible solution to the problems listed above and also allows to the scalability, process liability, adjusted testing even when a large number of students is being evaluated, creation of new opportunities for students (accessibility) and institutions (new markets) and development of methodologies that allow a secure and liable assessment in online distance courses.

This last item assumes vital importance for the University hence it's engaging in the development of online distance courses in lifelong learning area.

However the logistics required establishing CBA at a large scale in a university campus can be daunting. Rather unexpectedly there is lack of case reports on well succeeded programs. However the learned experience of coping with the problems of establishing an e-learning project in a large, spread and traditional university like University of Porto may be useful.

This project so far allowed us to identify strengths and weaknesses of computer-based assessment, with focus to automatic grading based on MCQ exams.

The group will continue research and implementation of good practices of new technologies for assessment.

Acknowledgments

The authors wish to thank Professor Daniel Moura, for his help and advice. Also to Professor Glória Queiroz wich provided the examination data and experience. The contribution from our teachers, faculties' technical personnel and students are warmly acknowledged.

References

- Ellis, R.A, Jarkey, N., Mahony, M.J., Peat, M. and Sheely, S. (2007) *“Managing Quality Improvement of eLearning in a Large, Campus-Based University,”* Quality Assurance in Education, vol. 15, pp.9-23
- Soeiro, A., Pinto, M.L. and Pereira, E. (2000) *Pedagogia e Novas Tecnologias: Estudos de Casos Universitários*, [CDROM], University of Porto Ed. (in Portuguese)
- Lawson, A.E. (1995) *Science Teaching and Development of Thinking*. Belmont, Calif.: Wadsworth, pp. 261-299.
- Lawson, A.E. (2003) *The Neurological Basis of Learning, Development and Discovery: Implications for Teaching Science and Mathematics*. Dordrecht, The Netherlands: Kluwer, pp. 225-260
- Gibson, E.J., Brewer, P.W., Dholakia, A., Vouk, M.A. and Bitzer, D.L. (1995) “A comparative analysis of Web-based testing and evaluation systems.”, Proceedings 4th WWW conference, Boston.

[Online References]

- Morten, P. and Keegan, D. (2006) *“Manchester Metropolitan University”*, Technical Report Megatrends in eLearning Provision NO/05/C/F/RF-83250, Norsk Kunnskaps-Institutt, Bekkestua, Norway, http://www.nettskolen.com/in_english/megatrends/MMU_article.pdf
- U.PORTO website, (2007), <http://www.up.pt>
- Falcão, R. (2006) *“E-learning at the University of Porto from 1998-2005: Implementing E-learning at a Traditional University”*, http://www.ai.tuwienac.at/elearning/innovation_lecture
- Senado da Universidade do Porto, Prémio Excelência e-LearningU.Porto, (2005), http://sigarra.up.pt/reitoria/LEGISLACAO_GERAL.ver_legislacao?p_nr=143 (in Portuguese)
- IMS Global Learning Consortium. IMS Question & Test Interoperability Specification (2006), <http://www.imsglobal.org/question/>
- Ferl Becta, *“On-line Assessment software”*,(2006), <http://ferl.bect.org.uk/display.cfm?page=193>
- S. Suleman, *“E-assessment”* (2003) Ferl First, BECTA, UK, <http://ferl.bect.org.uk/display.cfm?resid=5367&printable=1>
- Moodle, *Statistics*, (2006), <http://moodle.org/stats/>
- Moodle. *Release Notes*, (2006), http://docs.moodle.org/en/Release_Notes
- FastTest Pro. (2008).[Computer program]. Assessment System Corporation, <http://www.assess.com/FastTEST.html>
- Pearson Vue, *Become a test center* (2007), http://www.pearsonvue.com/pvtc/facility_reqs.pdf