Transition from traditional to electronic testing

Marjan Krašna
Faculty of Arts
University of Maribor
Koroškacesta 160,
2000 Maribor Slovenia
marjan.krasna@um.si

Tomaž Bratina
Faculty of Education
University of Maribor
Koroškacesta 160,
2000 Maribor Slovenia
tomaz.bratina@uni-mb.si

Branko Kaučič
Faculty of Education
University of Ljubljana
Kardeljevaploščad 16,
1000 Ljubljana, Slovenia
branko.kaucic@pef.uni-lj.si

Abstract. *In the past decade we have focused on production of e-learning materials. Different types of e-learning (distance learning, blended learning, m-learning ...) for younger generations needs to obey some didactical roles. We cannot expect that young generation would be self-regulated or self-controlled. Principles of programmed learning units with the control at the end have been successful in these situations for years. Evolution of implementation of ICT into education has come to the turning point. At the university level lecturers have no desire to produce better learning materials but they really like e-testing. From the first implementation of e-tests we have gradually learn the principles of e-test design. But the knowledge is not easily transferred to lecturers who used traditional testing methods. In the article we are going to present our experiences of e-testing that we solved in the past projects. From what we have learnt we prepare the recommendation for e-test preparation that are going to be used for in-house training at our faculties.*

Keywords:distance education, electronic testing, learning management system, knowledge assessment

1 Introduction

We have participated in the electronic learning development in our country. From the first attempts toward electronic learning around fifteen years ago we have active role in the progress. Now almost all higher level education institutions are involved into e-learning and now we have more than a decade of valuable experiences. A decade in ICT represents fairly long period of time. Every now and then old ideas are reinvented in new or more polished manner (i.e. Apple iBook application). Recently we showed that the common believe that Web 2.0 in "modern education"is not mandatory and some old fashioned (i.e. non-interactive) e-resources are highly regarded and frequently used by students [1]. What was once contemporary is not contemporary any more. It seems that in education field this time gap is even bigger. Partly this is due to amount of funds available for education but mostly due to the fear and the unwillingness to rush into newand educationally not yet proven technology. The latter is unfortunately also a common excuse not to use the technology that is proven and successful since users fear that when they start to use it, it will be outdated.

Conducting a digital augmented, blended, assisted, collaborative, etc. learning requires a significant amount of effort. Teachers (i.e. lecturers, assistants, tutors, etc.) should use contemporary approaches and technology, and prepare high quality visually appealing e-learning materials. In general, it is expected that they would teachthemselves how to produce them. But this is not true. Computer technology in education is mostly used for writing text and web browsing[2]. Clearly, changes in the education are not sufficiently quick to satisfy young generations (digi-kids or digital natives; generations Y and Z).

In the last few years we have established several learning management systems (LMSs) on our institutions. Additionally we have organized workshops about their usage, and teachers and students are now reasonably familiar with this technology. In general LMSs are now mostly used for learning material distribution, messaging and data warehouse for students' assignments. Recently several lecturers expressed their demands for additional training in LMS usage. Although we expected that teachers would like to upgrade their learning materials or use advanced types of communication, it was a surprise that their most desirable demand was how to efficiently use electronic testing (e-testing) within or outside their LMSs. In e-learning community e-testing has its history, benefits and also drawbacks. In order to prepare next workshops regarding the e-testing, we have researched the possibilities how to effectively use e-testing and help teachers to produce their own e-tests.

Authoring tools for traditional e-tests are available and easy to use. Traditional e-tests are also suitable for students since they can review answers before submission. LMS Moodle has even question bank and tests can be created randomly. In the article we are

going to present our evaluation of e-tests and suggestions for transition from traditional to e-tests.

2 Traditional testing

Assessing knowledge, competences and skills may be performed during different parts of learning process and for different purposes. Assessments can be performed for diagnostics reasons, commonly at the beginning of the learning process, in formative form during the learning process or in summative and formative form at the end of the learning process[3].Traditional assessment can be written (i.e. paper-and-pencil testing) or oral. Teachers are believed to be good evaluators for assessment of student's knowledge, competences and skills. Experienced teachers can use small amount of questions and based on received answers and observation of studentsbehavior can make reasonable accurate knowledge assessments. Sometimes oral examinationsare not feasible and written assignments are used, e.g. for large population of students. Teachers are still the best in the understanding and evaluationof written text answers, drawings and schematics answers, and can evaluate even partially correct answers.

To assess the level of knowledge, competences and skills desired testing instruments are required. They may differ regarding the purpose of the assessment. The assessment methods oriented exclusively to assess the level of knowledge are certainly different from those assessing competences and skills. But not always, since in pedagogical science and education the skills and knowledge are described as competences[4]. More precisely the competences can be defined as a combination of knowledge, understanding, skills, capabilities and values[5][6], or divided into knowledge, skills and wider competences[7]. In this text, only assessing the knowledge is taken into the consideration.

2.1 Important testing factors

For efficient and accuratetesting reliable instruments are required. Such instruments should meet certain requirements and/or characteristics.This is particularly important in grading answers. General and empirical characteristics are:reliability, discrimination, and difficulty. Rationalesand also important characteristics are validity, objectivity and sensitivity.

The *reliability index* expressed as the Chronbach's alpha (α) indicates the reliability of the instrument in measuring values of desired variable. By applying α to assessment of knowledge level, α represents the assessment's ability to interpret the level of knowledge being measured.The lowest acceptable value of Chronbach's alpha is 0.7 but in many cases lower values are also accepted[8].

In general, the assessment consistsof several questions which together influence the quality of the assessment. When the correct answer to the particular question is given by the majority of the best students in total assessment the question has an acceptable level of discrimination. The *level of discrimination* is expressed by using the point-biserial correlation coefficient[9]. The acceptable discrimination of the particular question is achieved by the value higher than 0.30[10].

The *difficulty index* of the assessment's particular question is the proportion of students answered the question correctly. The index is expressed as p-value percent (p%). The value between 50% and 80% is considered as the acceptable difficulty level of the particular question[11].

Validity of the assessment is the rational measure to determine if the assessment is measuring the desired characteristic. The validity can be verified by the comparison of the theoretical and operational definitions. More exact, it is the comparison of the achieved results with the results of similar assessments. To enhance the validity of the assessments the questions can be leveled according to widely adopted taxonomies like Bloom's taxonomy or other appropriate ones.*Objectivity* of the assessment is achieved if the assessment result is the consequence of the assessed characteristic without the influence of assessors. The *sensitivity* of the assessment is the characteristic that enables distinguishing small differences between the results which allows setting the minimal standards for students with lower level of knowledge[3].

When preparing the assessment the purpose and the goals should be defined first. Some experts recommend that questions should be prepared simultaneously during the learning process. The questions have to be clearly expressed, with unambiguous meaning, and enough attractive and interesting. The choices in all questions should be randomly sorted. The questions should be presented in ascending difficulty order.

In traditional testing two main types of questions are used: filling-in type and short answers type. The latter can be implemented as three different versions: as single or multiple choices, alternative questions and matching items[12].

All mentioned rules, recommendations and examples in this Section are applicable in e-testing, with or without some reasonable modifications. Some of them may be omitted in future development of e-testing and certainly some new principles will arise.

3 Computer based testing

Computer based testing (CBT)[13]and technology-based assessment (TBA) is fast. Students get feedback as soon as they finish the test (i.e. submit all answers to the system). It has many types of questions and can be used in most of situations. Traditional e-tests enable students to answer questions at random order until they submit results or run out of time.

3.1 Types of e-tests

There are two major types of electronic tests: software programs or browser based tests. If we create software program then we have better control over the test subject and more custom made test questions. On the other hand if we prepare browser based tests then we are limited with the browsers capability to display graphical elements and manipulate with them (Table 1).

Table 1: Advantages and disadvantages of e-tests

	Advantages	Disadvantages
Program based tests	• Ability to prepare non-traditional test questions • Can be used offline • Better control over user actions	• Require specific software platform • Custom build tests • Closed tests with limited ability to modification
Browsers based tests	• Works on any operation system • Many authoring tools are readily available • Easy to learn	• Different browser behavior and capability • Sufficient security cannot be achieved • Reliability depends on internet connection

3.2 Types of questions for e-tests

There are many authoring tool to prepare electronic tests. From freeware to commercial software all have common types of questions: Yes/No (True/False) questions; Multiple choices; Matching; Short answer; Numerical; and Text or Essay (Table 2).

Some authoring tools may have additional questions types that can be described as subtypes of these general types. If we take a look at tests in LMS Moodle we see additional types as Description, Embedded answers (Cloze), and Random Short-Answer Matching.

Table 2: Advantages and disadvantages of different question types

	Advantages	Disadvantages
Yes/No (True/False) questions	• Simple questions • Question and statements covered. • Can be used in different way (positive, negative and combined)	• Question design • Tests with only this type of question are hard
Multiple choice	• Use in different way (one right answer, multiple correct answers) • Most common type of questions	• Require additional design consideration since not all answers can have the same weight • It is not always easy to prepare the answers that are not trivial to dismiss
Matching	• Dynamic questions • Require more than just clicking • Can be used for matching or categorizing	• Not all authoring tool enables the classification (multiple answers to put into the same category) • Not all browsers are capable to handle moving objects
Short answer	• Appropriate in language learning	• Not all authoring tools enable multiple correct answers. • Sometimes correct answer can be lost.
Numerical	• More suitable for numerical answer then text • Type checking	• Sometimes when the result is real number true answers can be lost because of rounding
Essay	• Enables creativity of student	• Cannot be electronically evaluated • Post processing is required • Response is delayed • Due to technical limitation sometimes special characters can corrupt data

Traditionally we usually prepare test that is used for all students. But in the electronic testing we have ability to randomly prepare tests and each student can have its unique test. Randomly created unique test can only be done in LMS environment and not all authoring tools are able to copewith the specifications.

LMS Moodle support basic question types and can incorporate tests from other authoring tools. One difference we have discovered in Moodle is the behavior of "multiple choice" when used with multiple answers. These types of questions behave different in different authoring tool. In some authoring tools all correct answers need to be checked to receive points but on other (Moodle included) each correct answer brings points. Both behaviors can be treated as correct depending on situation. For example 1 (Figure 1) a student should receive points for any right answer he/she pick but for example 2 (Figure 2) only all correct answers enable him/her to finish the job. In these types of questions we also need to give penalty factors for wrong answers and gross sum of all points if all were checked should be 0. However, there are still debates about this and some said that each correct answer could give less positive points than false answer gives negative points. For example correct answers score 1 point but wrong answer scores -2 points. This would prevent students to answer questions at random since they would know for sure they could not pass.

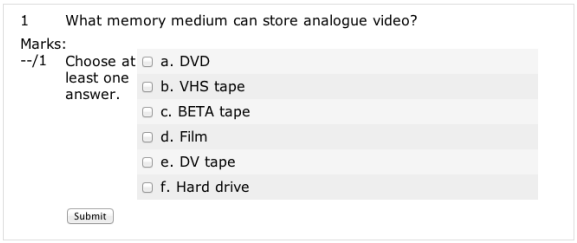

Figure 1. Multiple choices question that can be partially correct since answers are not related.

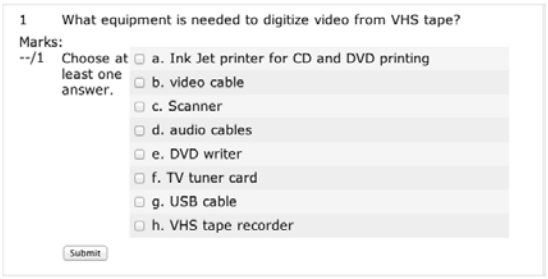

Figure 2. Multiple choice questions that need all correct answers because digitalization cannot be achieved if one answers is missing

4 Suitability of tests

It is evident that all tests are not equally suitable in every situation. From the Figure 3 we see different types of tests and their complexity levels. With the raise of complexity there are also raise of potential problems. If we want to examine the level of knowledge we can use positive or negative question type. Positive is usually simpler than negative and the question arouse. How can we explain the situation where students would answer positive question type and not negative? Do we grade knowledge or do we grade understanding of question?

	Complexity					
Openness	Action	I	II	III	IV	V
	Select	True / False	Alternate choice	Multiple choice	Implicit answers	Certainty degree
	Identify	Multiple True False	Yes/No with explanation	Multiple Answer	MA with images	Click on image zones
	Match	Matching	Categorizing	Sequencing	Prioritizing	Assembling proof
	Correct	Remove intruder from a list	Jumbled sentence	Detect wrong form filling	Click errors on an image	Problem solving
	Complete	Fill-in blanks	Drop down fill-in blanks	Fill-in form	Calculated answer	Listening comprehension
	Construct	Lab simulation	Analysed Open Answer	Link concepts on a map	Draw valid sequence	Delineate zone on an image
	Project	Open answer / Essay	Word assignment	Spreadsheet assignment	Presentation assignment	Multimedia project
	Collaborate	Forum discussion	Documents sharing + peer review	Group publication	Team blog with roles	Problem solving in team

Figure 3. Complexity and openness in tests/assignments of knowledge assessment [14].

From the Figure 3 we know that inside category increased complexity means harder questions but we cannot know which type of question is harder between categories (i.e. Identify V or Correct IV). From our experiences we know that the second on is much harder. Sequencing is not easy but is still easier than fill-in blanks.

5 Changing the way we think for successful e-tests

It may seem easy to design e-tests but obviously this is not true. In the projects of e-learning materials production we were involved as authors and technical support. Teachers often give us insufficient data to implement e-tests and this produce mutual frustration. As if we did not speak the same language. Our examples how they should prepare tests were different then their traditional tests preparation. After many tries and errors we discover the right approach. ***Teachers should prepare tests for someone else to evaluate them.*** With this in mind we were able to cooperate more fluently but some faults were still detected in communication.

To give a better overview about the misunderstandings we will give some examples.

5.1 Yes/No question type example

Teachers give us the question: Can we catch the air? We have to ask him what he expects the students would answer for this question and what is the right answer. He responded: Students will answer this question you should just put it on the screen. We have changed another couple of e-mails for him to understand what we need and results are seen on Figure 4.

5.2 Click on image zone example

In the primary school students should be able to understand the concept of data in tables. On the Figure 6 we see school time schedule. For e-test teacher gives us the question: Where you have mathematics? Question would be totally suitable for class environment but not for e-learning. It raised a problem of understanding. What happens if student have mathematics in his own schedule somewhere else than it see it on presented table. After negotiation we redesign the question to: What is on the schedule on second hour on Wednesday? This question is universally understood and cannot be misinterpreted.

5.3 Matching question type example

Teacher send us Word document where in two columns of different words. When we start preparing matching question we find out that horizontally words to not match association. This may be huge problem if matching were on the expertise level where we could not notice the fault. When we asked the teacher to show us the right answers she replayed that student will solve it and we should not worry about the test. Even when we said to her that we need right answer she replied that if all answers were evident then all items would be horizontally connected and the

question would be stupid. We have to show her that computer needs solution and students will receive shuffled responses on two different examples of matching question for her to understand the concept (Figure 13). After few weeks we noticed that she forgets the concept and needs knowledge refreshment.

6 Examples of e-tests in our projects

In this topic we present the question types of different complexity level for different learning groups. Questions' design process should consider students'didactically suitability and comprehension level.

6.1 Primary school tests

In the lower levels of primary schools we have to mask tests into nice looking computer screens. In this section we show types of questions we used for evaluation of knowledge retention from e-learning materials. We have used basic question types (yes/no typeFigure 4, multiple choicesFigure 5, click on imageFigure 6) and one composite question type (sequence and click on image zoneFigure 7).

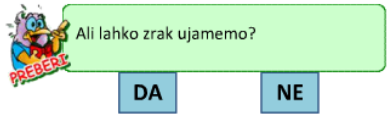

Figure 4. Yes/No test.

Figure 5. Multiple choices

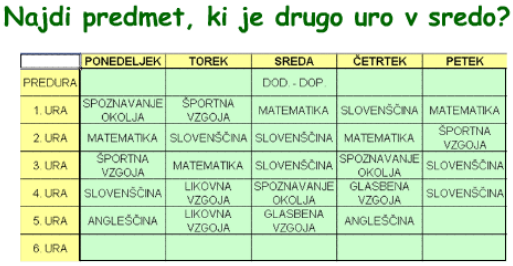

Figure 6. Click on image zone

Figure 7. Sequence and click on image zone

6.2 Secondary school test

Secondary school test can include all question type that are available in authoring tools. Most common types are multiple choices (Figure 8), multiple answers (Figure 9) and click on image zone (Figure 10) question types. With the question type fill-in blanks (Figure 11) we need to be careful if question can have more than one answer. Computer is not realy good in evaluation of such types of question and designer shoud know potential problems.

Figure 8. Multiple choices

Figure 9. Multiple answers

Figure 10. Click on image zone

Figure 11. Fill-in blanks

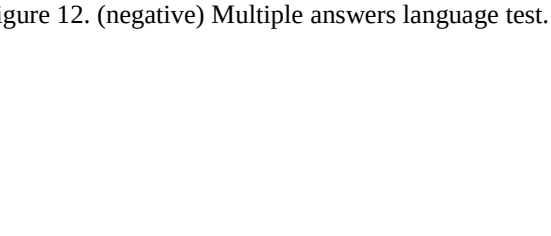

Figure 12. (negative) Multiple answers language test.

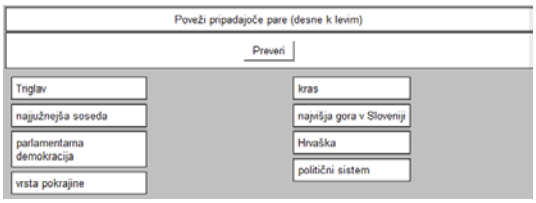

Figure 13. Matching language test.

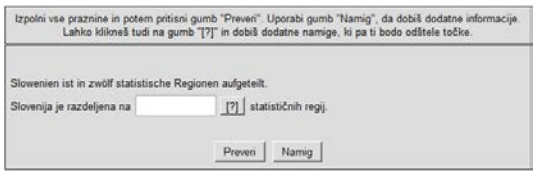

Figure 14. Fill-in blanks language test.

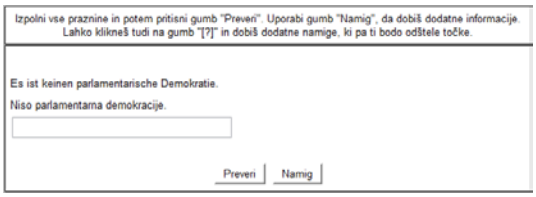

Figure 15. Fill-in whole sentence language test (extremely difficult).

6.3 Language understanding tests

In the projects for Slovene language for foreigners designers used different question types. From traditional yes/no question types, (positive) multiple choices, (negative) multiple answers (Figure 12), to matching (Figure 13) and fill-in blanks (Figure 13). In our technical opinion the fill-in blanks (Figure 15) are the most difficult question types since no one was able to answers all of the questions correctly.

7 Conclusion

The development is never strait forward process. Even if we are aware that e-learning materials still far from perfection the most demands from teachers are not their improvements. From the in house training we find out that most desirable topics in e-learning are e-tests. Reasons are clear and expected but we fear that rush may not provide satisfactory results. E-tests require a mental change from traditional tests. Not all teachers have ability to switch immediately and some training and support will be required. Persons responsible for transition from traditional test to electronic test must highlight the problems of e-tests and restrain teachers to use the same concept from traditional test (fill-in blanks) too often. The poor results may increase frustration of students beyond acceptable level.

References

[1] R. Repnik, M. Krašna and B. Kaučič, "E-learning in the modern curriculum development," in *E-learning / book 5*, Rijeka, InTech, 2012, p. (In print).

[2] I. Gerlič, T. Bratina, N. Veček and T. Pungartnik, "Stanje in trendi uporabe informacijsko komunikacijske tehnologije (IKT) v slovenskih osnovnih šolah," University of Maribor, Faculty of Natural Sciences and Mathematics, Maribor, 2011.

[3] B. Marentič Požarnik, Psihologija učenja in pouka, Ljubljana: DZS, 2000.

[4] EU legislation, „Key competences for life-long learning,“ [Elektronski]. Available: http://europa.eu/legislation_summaries/education_training_youth. [Poskus dostopa 2010].

[5] J. González in R. Wagenaar, „Universities' contribution to Bologna process. Project Socrates,“ University of Groningen, Bilbao and Groningen, 2005.

[6] University of Maribor, Faculty of Natural Sciences and Mathematics, „Development of Natural Science Competences,“ University of Maribor, Faculty of Natural Sciences and Mathematics, Maribor.

[7] S. Avsec and B. Kaučič, "Competence modelling on solid waste management," in *Competence modelling for vocational education and training : innovations for learning and development*, 2011, pp. 71-81.

[8] A. J. Reynaldo, "Cronbach's Alpha: a tool for assessing the reliability of scales," *Journal of Extension*, vol. 37, no. 2, 1999.

[9] J. Sagadin, Poglavja iz metodologije pedagoškega raziskovanja 2. izd., Ljubljana: Zavod Republike Slovenije za šolstvo in šport, 1993.

[10] D. Žagar, "Napotki za pripravo pisnih preizkusov znanja v devetletni osnovni šoli," *Vzgoja in izobraževanje*, vol. 23, pp. 18-21, 2002.

[11] M. Ivanuš Grmek, B. Čagran and L. Sadek, Eksperimentalna študija primera pri pouku spoznavanja okolja, Ljubljana: Pedagoški inštitut, 2009.

[12] University of Maribor, Faculty of Education, "Testi znanja: prednosti in pomanjkljivosti," University of Maribor, Faculty of Education, Maribor, 2009.

[13] C. G. Parshall, J. A. Spray, J. C. Kalohn and T. Davey, Considerations in Computer-Based Testing, New York: Springer, 2002.

[14] T. De Praetere, "E-Learning: Learning throught the use of devices," n.d.. [Online]. Available: http://knol.google.com/k/e-learning#. [Accessed 17 01 2012].

[15] Ministry of Education and Sport, Republic of Slovenia, „Republika Slovenija, Ministrstvo za šolstvo in šport,“ Januar 2009. [Elektronski]. Available: http://www.mss.gov.si/si/okroznice_razpisi_in_javna_narocila/javni_razpisi/?tx_t3javnirazpis_pi1%5Bshow_single%5D=941.