

Past, Present and Future of e-Assessment: Towards a Flexible e-Assessment System

Mohammad AL-Smadi, Christian Gütl

Institute for Information Systems and New Media (IICM), TU-Graz, Austria

Key words: *Assessment, E-Assessment, CBA/CAA, E-Learning, Instructional Design*

Abstract:

The rapid change in our entire life influences directly or indirectly the systems that controls our knowledge, skills and behaviour. The evolution of our culture is one of the major indicators for this change. Our educational system has been influenced by this rapid change over the time and technology is increasingly used in learning settings. Assessment as a part of the educational system is exposed the same changes. Specialists have been taken care of adapting the learning system with culture changes but unfortunately they have not properly focused on performance measures and feedback. Assessment research trying to keep pace with modern learning settings but still provides room for interesting and challenging research. In this paper, we give an overview about e-assessment history, discuss e-assessment rationales, and outline the main challenges of using computers to assist the assessment process. Finally, we introduce a concept for a flexible e-assessment system applicable in a great variety of learning settings.

Introduction

Using computers to assist assessment task has been an interesting research topic for decades, however, developments have mainly transferred traditional assessment approaches into computer environments. Moreover, in order to automatically grading students' assignments, types of assessment approaches have been further limited. [15] Consequently, the rapid increase of using technology in learning settings expedites also the need for new technology-based assessment. Our life has been influenced by a revolution in the field of information and technology. As a result, peoples' mentality has changed significantly in the recent years. Consequently, pedagogy has become affected and educationalists have also started redesigning educational systems [33]. Learning is no more divided; there is no separation between schools' education and workplace experience. Acquiring knowledge is a continuous learning process. According to [25], Learning is a continuous process over lifetime, it is a lifelong process. Therefore a new paradigm for assessment in lifelong learning is becoming important. Changing education from memorizing facts to higher levels of comprehension and synthesis requires building and assessing critical-thinking skills. According to [19], measuring knowledge is important but is not enough. The academic programs should work on building and assessing students' critical-thinking skills.

In general, assessment has different strategies according to its purposes. The two main basic types of these strategies are formative and summative assessment. *Formative assessment* is part of the learning process; this assessment is used to give feedback to both students and teachers in order to guide their efforts toward achieving the goals of the learning process. Where, *summative assessment* is performed at the end of specific learning activity; and used to judge the students progression and also to discriminate between them [5]. According to [2], technology is an essential component of modern learning system. As a result, technology is also increasingly needed for the assessment process to be authentic.

E-assessment can be distinguished as *Computer Based Assessment (CBA)* and *Computer Assisted Assessment (CAA)* which are often used interchangeably and somewhat inconsistently. CBA can be understood as the interaction between the student and computer during the assessment process. In such assessment, the test delivery and feedback provision is done by the computer. Where CAA is more general, it covers the whole process of assessment involving test marking, analysis and reporting [9]. The assessment lifecycle includes the following tasks: planning, discussion, consensus building, reflection, measuring, analyzing, and improving based on the data and artifacts gathered about a learning objective [30].

The type of useful assessment method depends on the learning objectives. These objectives have been classified in Bloom's Taxonomy into six levels: knowledge, comprehension, application, analysis, synthesis and evaluation [4] [26]. Consequently, a variety of exercises which assess the different objectives' levels should be applied. E-assessment systems can be classified according to the nature of the users' response to the test items into, *fixed response* systems and *free response* systems [11]. According to [11], fixed response systems which also referred to as *objective* forces the user to have a fixed response by selecting an answer from a pre-prepared list of solution alternatives. Where, in the free response systems *non-objective*, unanticipated answers formulate the user's response. In such type of systems skills like programming, essays writing and meta skills are assessed rather than fact or knowledge assessment which represents the main domain of the first type. Additionally, portfolios can also be used to assess learning outcomes. Moreover, according to [10], portfolios represent the highest point of students' learning, what they collect, assemble and reflect on samples are represented in their portfolios.

E-assessment is not only applicable for individuals, but it is also used for groups. Assessment of groups, also referred to collaborative assessment, is used to assess the participation of individuals in group work and their behavior of how they collaborate with each other to solve problems [35].

In this paper we explore how assessment has developed over time, and based on the most important challenges identified a vision for a flexible e-assessment system is introduced. To this end, the remainder of the paper is organized as follows: an overview of the historical evolution of e-assessment is given in section 2. The rationales and the motivations for e-assessment systems are explored in section 3, followed by the challenges that any e-assessment system may face or the ones that are posed by the use of e-assessment during learning in section 4. Based on that, flexible overall e-assessment architecture is discussed in section 5. Finally, section 6 concludes with insights and intended future work.

Historical Overview of E-Assessment

Computers have been used for decades to assist assessment. One of the earliest attempts of using computers to assist instructional and assessment process refers to the early 1960's when PLATO (Programmed Logic for Automatic Teaching Operations) project was started at the University of Illinois [40]. TICCIT (Time-Shared, Interactive, Computer-Controlled, Information Television) [20] which has been started in 1967 is another example of a large-scale project for using computers in education. The history of e-assessment can also refer to the use of computers to automatically assess the students' programming assignments [14]. One of the early attempts of using computers to automate the process of assessing students' programming assignments was the "Automatic Grader" [21]. Rather than using this program as a compiler for the programming assignments, it also helped the student to better learn programming, and also facilitates the teacher to supervise a larger number of students at the same course. Another application for the Automatic Grader was long distance teaching [21]. Authors of [16] presented another system for automatically assessing programming exercises written in Algol. The system was used by the students of a numerical analysis course at the

University of Stanford to assess their programming exercises. The system was responsible of data supplying, running time monitoring and keeping a “grad book” for recording problems.

Assessment played a main role for enhancing the performance of learners and also the quality of instructional materials. According to [34], in 1960’s formative evaluation was applied to the drafts of instructional materials before they were in their final form. Assessment as main part of the instructional design and media was affected by the revolution of micro-computers in the 1980’s [34]. According to Reiser, during the 1980’s and afterwards an increasing interest of using computers in instruction had started and computers were used in automating some instructional design tasks. Assessment Systems in other fields such as mathematics [37] and chemistry [31] appeared after soon.

The 1990’s was affected by the important impact of the World Wide Web (WWW). Since then assessment systems started to be web-based systems. Blackboard.com [3] provides automatic grading of multiple choice and True/False questions. Systems such as QUIZIT [38], WebCT [39], ASSYST [24] and PILOT [6] are also examples of web-based systems with ability of online testing and grading. Recent examples such as, [28] presents a web-based assessment system that applies Bloom’s taxonomy to evaluate the outcomes of students and the instructional practices in the educators in real time. In a step towards a fully automatic knowledge assessment, Guetl in [18] introduced the e-Examiner as a tool to support the assessment process by automatically generating test items for open-ended responses, marking students’ short freetext answers and providing feedback.

E-Assessment Motivations and Rationales

Motivations and rationales of using e-assessment instead of paper-based assessment in higher education are discussed in this section. According to Charman and Elms [9], the practical and pedagogic rationales are the main motivators for adopting e-assessment in higher education.

Practical Rationales

Increasing number of students supervised by the same staff resources causes an increase in the staff workload. Accordingly, time spent by the teachers to assess students is also increasing. Therefore, a step toward the e-solutions becomes a real need. Although many e-learning environments have been developed in universities to overcome the workload problem, most system have not adequately solved the assessment tasks. Therefore, reducing time and efforts spent on students’ assessment is a strong rationale to use the e-assessment technology [9].

Pedagogic Rationales

Formative assessment is an important component of any learning systems. In such type of assessment, the current state of students is diagnosed. This ongoing process involves a continuous feedback to both teachers and students which allows them to enhance their teaching and learning activities to satisfy the learning goals [5]. There is an increasing pressure on teachers in further and higher education to provide assessments that are fair, reliable, efficient and effective. Brown and Race in [7] illustrated that each assessment must take proper attention of a set of quality dimensions. Many of the values are achieved by CBA systems in nature (*Fair*: offer fair opportunity for success; *Equitable*: be indiscriminating between students; *Formative*: give many opportunities to learn through feedback; *Well timed*: provide learning stimulus and be fair; *Redeemable*: allow a series of opportunities; and *Efficient*: be manageable within the constraints of resources); some others (*Valid*: accurately assess the delivered material, *Reliable*: promote consistency between assessment tasks, *Incremental*: increase reliability and consistency over time; and *Demanding*: challenge students and ensure high standards) depend on the experience of the assessment designer or the system designer [9].

According to [12], appropriate assessment information provides an accurate measure of student performance to enable students, teachers, administrators and other key stake holders to make effective decisions. Therefore, any CBA/CAA system should satisfy the quality dimensions outlined above.

E-Assessment Challenges in Modern Learning Settings

Digital Mimes

It has been a long time since the society started thinking of transferring assessment to computer-based environments. Since 1960's several steps towards achieving this goal have been taken. But unfortunately, e-assessment has been criticized for imitating the conventional assessment. The author in [15] argues that e-assessment system's designers imitate traditional assessments. He also stresses that these systems only support limited number of exercises types. E-assessment is also criticized that it still assumes that students have to retain the context related information in their memories. To overcome this problem, Ridgway et al in [36] suggested *open-web examinations* which are similar to the open-book examinations and can lead to desirable learning. From a more general viewpoint, according to [30], the assessment process begins by identifying the learning goals and objectives and must also consider the characteristics of the learning setting. Consequently, e-assessment systems must follow these guidelines which also include to properly use available functionality from technology.

Mentality Change & Culture evolution

The rapid change in our society culture and the increasingly use of technology in our modern life activities influences directly our educational process. Students today are considered to be multi-tasked; they grown up with technology around them all over their lives, they use technology anytime, anywhere. Prensky termed them as Digital Natives [33], where he also identified a gap between students and our educational systems. The current applied educational systems are not appropriate anymore. Prensky also discusses the gap between the students (Digital Natives) and their instructors who are not skilled enough to use the technology (Digital Immigrants). As a result, students who applying new technologies in new ways also need new assessment mechanisms based on these technologies and skills. Therefore, a modernizing step is also required for our educational assessments which will be influenced by the abilities of Web 2.0 [15].

Security & Privacy

One of the major challenges of e-assessment is security. The problem of having a secure, fair and effective e-assessment system is not new. The same problem has been a major concern of e-learning systems in general. Several solutions have been applied to avoid this problem over the years. Marias et al in [29], classify the security of e-assessment into two main types. The first is Web security: this type of security concerns to the security of servers where the web application is running. The second is e-assessment security: which concerns the authenticity of users and supervision of assessment location. Several solutions have been discussed by the authors varied from the simplest cheapest way of using passwords to the more complex and more expensive techniques of adding physical instruments to the assessment system such as, Biometric authentication and using video conferencing setups [1]. Maintaining users' privacy is another concern, which aims to protect specific users' data not to be accessed by other users. A big issue in this context is to find a good compromise of providing adequate information of students' performance to teachers and keeping specific details secret [8].

Assessment and Feedback as a Mean to Learn

Assessment is not only used for measuring and judging students work. Furthermore, it is a mean for learning, where different forms of assessment such as portfolio-assessment, self-assessment and peer-assessment are used to encourage students to further progress and learn [13]. For example web-based exercises require students to search for solutions where they actually are learning. Most of these forms of assessment are formative. Feedback as a spirit of formative assessment is also a mean by which the objectives of using such type of assessment are achieved. Feedback has been considered to be a mirror for learning [26]. Both learners and educators can use this mirror to see what they have done and also what they have not during the course. Therefore, we believe that feedback is an essential mean for achieving the objectives of using formative assessment. Technology has a major influence on feedback [9]. The use of computers facilitates the process of tracking user behavior and performing assessments as well as analyzing the results. Consequently, it is much easier to provide feedback about students work immediately or timely during courses. As a result, they can easily find their ways to success.

Interoperability & Standards

Interoperability is defined in [8] as the ability of different systems to share information and services in a common format. To widely use a flexible assessment tool, information such as questions/exercises and answers, users' information, list of enrolled students, courses information and learning objectives must be shared with other systems and tools. Several standards such as IMS-QTI (IMS Question and Test Interoperability) [22], PAPI (Public and Private Information for Learner - IEEE) [32], GESTALT (Getting Educational Systems Talking Across Leading-edge Technologies) [17] and IMS LIP (IMS Learner Information Package) [23] have been proposed for such purposes. Therefore, e-assessment systems should be designed to be flexible and deal with most of these standards. Consequently, it will be able to communicate and interact with other systems in an open way.

E-assessment Automation & Assessment Types

As mentioned earlier, the assessment process begins by identifying the learning goals and objectives [30]. As a result, a variety of assessment types (such as limited choice exercises, open ended questions and essays) are used to achieve these objectives. Most of the developed e-assessment tools are related to specific part(s) of the assessment cycle or limited to some type(s) of assessment. Unlike those tools we believe that e-assessment tools should support the whole cycle of assessment and to be designed based on the learning goals. Despite the difficulties of automatic questions generation, automatic marking and grading or even automatic feedback provision, we advocate that e-assessment tools should support the entire lifecycle of assessment by (semi-)automatic methods. A comprehensive literature survey and a first approach towards an automatic assessment tool can be found in [18].

A Proposal for an Enhanced E-assessment System

Based on the discussion in previous sections, we are proposing a flexible e-assessment system which includes: (a) *flexible design* to be used as a stand-alone service or to be easily integrated in existing systems. (b) *User-friendly interfaces* for both students and educators where a user interaction and online submission of solution and evaluation can be done. (c) Assessment environment for *various learning and assessment settings* which supports guided as well as self-directed learning. (d) *Management and (semi-)automatic support* over the entire assessment lifecycle (exercises creation, storage and compilation for assessments, as well as assessment performance, grading and feedback provision). (e) *Rubrics design* and implementation interfaces to allow the educators to design their own rubrics based on learning

objectives to assess learners' performance against a set of criteria. (f) *Support of various educational objectives and subjects* by using various tools sets which for example enables automatic exercise generation or selection, automatic grading and feedback provision. (g) *Results analysis and feedback provision* (immediately or timely) of the current state of user knowledge and metacognitive skills for both educators and learners and also for adapting course activities and learning contents based on users' models. (h) *Standard-conform information and services* to be easily sharable, reusable and exchangeable. This will include the tests' questions, answers and students' results, rather than any other required services. And finally, (i) *Security and privacy* where a secure logon of users based on pre-defined levels of access, and also users' authentication based on machine (domain users) or by usernames/passwords.

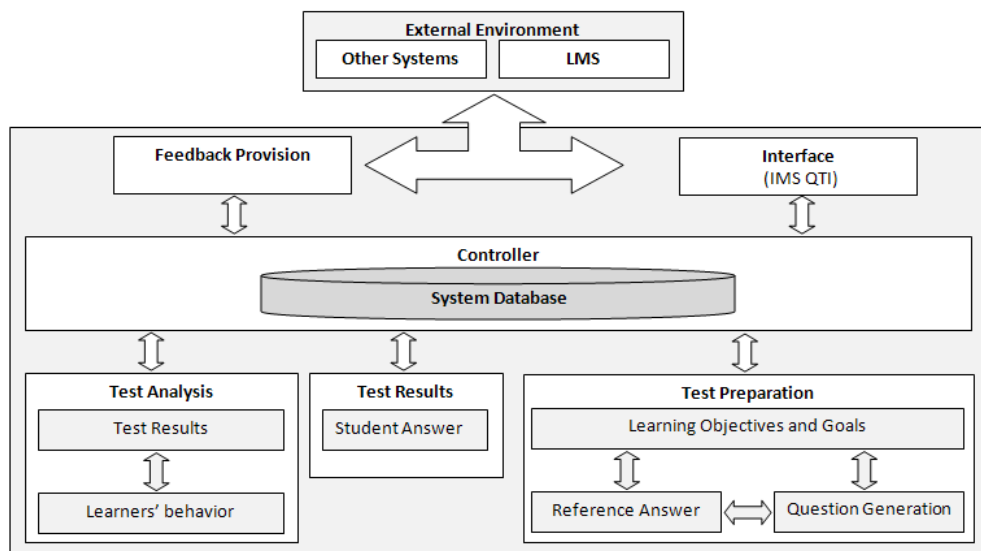


Figure 1. Overall Conceptual Architecture

As depicted in Figure 1, a first attempt for an overall conceptual architecture of the proposed system consists of a set of composable tools which represents the functionalities for a specific assessment setting. The controller module is the core module where the whole assessment lifecycle is managed as well as the communication between the other modules is handled. The test preparation module is responsible for both (semi-)automatically questions generation and reference answers preparing. The automatic questions generation can be done by extracting questions from the courses' content based on specific objectives and goals, the module can also assist the educators to design their exercises using user-friendly interfaces designed for this purpose. In both cases, questions will be generated in regard to common standards (such as IMS QTI) to maintain the ability of interoperability discussed earlier. The test results module assesses students' answers regarding to pre-defined rubrics which are used to assess learners' performance against a scoring scale and specific criteria. The test analysis module analyzes the output of the test results module and the records of learners' behavior and provides feedback through the feedback provision module to the external environment. The interface module handles the information flow between the tool and the external environment (Learning Management Systems (LMS) or any other standalone assessment tools or systems) in a standardized manner.

Conclusion and future Work

In this paper we have outlined the increasing need of e-assessment for modern environments in educational settings. We have shown how technology has been a boost to evolve

assessment process during several time periods. E-assessment history has been explored and important e-assessment systems have been exemplarily mentioned. The rationales and the motivations for using technology in assessment have been discussed, where we have classified the rationales for the e-assessment as practical and pedagogical ones. We have also outlined the possible challenges that people may face while they are designing or introducing e-assessment systems.

The need to innovate and not imitate the conventional assessment is one of the main challenges. Multi-tasked learners with traditional and limited educational systems are senseless. Other challenges related to the design and implementation of the e-assessment system such as security, privacy, feedback and standards have also been discussed. As a result of the broad review of the topic e-assessment, we have introduced requirements on an abstract level for a flexible e-assessment system and have shown a first attempt to an overall conceptual architecture.

Based on the first conceptual architecture, we will start an in-deep requirements analysis and will start the first phase of a prototype implementation.

References:

- [1] Barker, T. and Lee, S. Y. (2007). The verification of identity in online assessment: A comparison of methods, Proceedings of 11th Computer Aided Assessment Conference, Loughborough, July, 2007.
- [2] Bennett, R. E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *Journal of Technology, Learning, and Assessment*, 1 (1).
- [3] Blackboard Inc. Retrieved on 7 July 2008 from www.blackboard.com.
- [4] Bloom, B.S. (1956). *Taxonomy of educational objectives, Handbook I: the cognitive domain*. David McKay Co Inc., New York.
- [5] Bransford, J.D., Brown, A.L., Cocking, R.R. (Eds.) (2000). *How People Learn: Brain, Mind, Experience, and School. Expanded Edition*. Washington DC: National Academies Press.
- [6] Bridgeman, S., Goodrich, M.T., Kobourov, S.G., Tamassia, R., (2000). PILOT: An Interactive Tool for Learning and Grading, In Proc. of SIGCSE 3/00 Austin, TX, USA, ACM Press, pp. 139-143.
- [7] Brown, S., Race, P. (1996). *500 Tips on assessment*, Cogan Page, London, UK.
- [8] Bull, J. & McKenna, C. (2001). *Blueprint for Computer Assisted Assessment*. London: RoutledgeFalmer.
- [9] Charman, D., Elms, A. (1998). *Computer Based Assessment: A guide to good practice, Volume I*, University of Plymouth.
- [10] Chun, M. (2002). Looking where the light is better: A review of the literature on assessing higher education quality. *Peer Review. Winter/ Spring*.
- [11] Culwin F.,(1998). Web hosted assessment: possibilities and policy, Proceedings of the 6th annual Conference on the Teaching of Computing/3rd Annual ITiCSE Conference on Changing the Delivery of Computer Science Education, Pages 55–58.
- [12] Dietal, R. J., Herman, J. L., & Knuth, R. A. (1991). What does research say about assessment?, *North Central Regional Educational Laboratory*. Retrieved on 14 April 2008 from: <http://methodenpool.unikoeln.de/portfolio/What%20Does%20Research%20Say%20About%20Assessment.htm>.
- [13] Dochy, F.J.R.C.. & McDowell, L. (1997). Introduction: Assessment as a tool for learning, *Studies in Educational Evaluation*, 23 (4), 279-298.
- [14] Douce C., Livingstone D., and Orwell J., (2005). Automatic test-based assessment of programming: A review, *Journal on Educational Resources in Computing*, ACM.
- [15] Elliot B., (2008). Assessment 2.0: Modernising assessment in the age of web 2.0. Scottish Qualifications Authority, Retrieved on 28 May 2008 from: <http://www.scribd.com>.
- [16] Forsythe G. E., and Wirth N., (1965). Automatic grading programs, *Communications of the ACM*, 8(5):275-278.
- [17] Gestalt (1999). *GESTALT: Getting Educational Systems Talking Across Leading-Edge Technologies; Work Package 3, D0301 Design and Specification of the RDS (Resource Discovery Service)*, 1999. Retrieved on 7 July 2008 from <http://www.fdgroupp.co.uk/gestalt>.

- [18] Guetl C. (2007). Moving towards a Fully Automatic Knowledge Assessment Tool, *extended paper of the IMCL 2007 paper, iJET International Journal of Engineering Technologies in Learning*.
- [19] Haken, M. (2006). Closing the loop-learning from assessment. Presentation made at the *University of Maryland Eastern Shore Assessment Workshop*. Princess Anne: MD.
- [20] Hayes, B. G. (1999). Where's the data? Is multimedia instruction effective in training counselors? *Journal of Technology in Counseling* 1.1 [On-line]. Retrieved on 7 July 2008 from: http://jtc.colstate.edu/vol1_1/multimedia.htm.
- [21] Hollingsworth, J. (1960). Automatic graders for programming classes. *Commun. ACM*, 3(10):528–529.
- [22] IMS (2005). IMS Question and Test Interoperability Overview. Version 2.0 Final Specification, Unpublished, 2005, Retrieved 19 June 2008 from: http://www.imsglobal.org/question/qti_v2p0/imsqti_mdudv2p0.html.
- [23] IMS LIP (2005). IMS Learner Information Package (IMS LIP); Version 1.0.1 IMS Learner Information Package Specification, IMS Global Learning Consortium Inc., (January,17), Retrieved on 7 July 2008 from: <http://www.imsglobal.org/profiles/index.html>, last visit: 2004-11-05.
- [24] Jackson, D., and Usher, M. (1997). Grading student programs using ASSYST. In *Proc. 28th SIGCSE Tech.Syrup.*, pp. 335--339.
- [25] Jegede O. (2005). Towards a New Paradigm for Assessment in Lifelong Learning, *Presented at the International Association for Educational Assessment (IAEA) Conference* held at Abuja from 5-9 September 2005.
- [26] Krathwohl, D.R., Bloom, B.S., and Bertram, B.M., (1973). Taxonomy of educational objectives, the classification of educational goals. Handbook II: affective domain. David McKay Co. Inc., New York.
- [27] Klassen, J (2001). Pedagogical Support for the use of Information technology in teaching, *Conference Proceedings for Informing Science*, Krakow, Poland.
- [28] Lei He. (2006). A novel web-based educational assessment system with Bloom's taxonomy. *Current Developments in Technology-Assisted Education, Published by FORMATEX, Badajoz, Spain*, Volume III, 1861-1865.
- [29] Marais, E., Argles, D. and von Solms, B. (2006). Security Issues Specific to e-Assessments. In: *8th Annual Conference on WWW Applications*, 6-8th September, 2006, Bloemfontein.
- [30] Martell, K., & Calderon, T. (2005). Assessment of student learning in business schools: What it is, where we are, and where we need to go next. In K. Martell & T. Calderon, *Assessment of student learning in business schools: Best practices each step of the way* (Vol. 1, No. 1, pp. 1-22). Tallahassee, Florida: Association for Institutional Research.
- [31] Myers R. (1986). Computerized Grading of Freshman Chemistry Laboratory Experiments, *Journal of Chemical Education*, Volume 63, Pages 507-509.
- [32] PAPI (2000). Public and Private Information; IEEE 2000, Draft Standard for Learning Technology - Public and Private Information (PAPI) for Learner, IEEE P1484.2/D6, 2000. Retrieved on 7 July 2008 from: <http://ieeeltsc.org>
- [33] Prensky, M. (2001). Digital Natives, Digital Immigrants. In *On the Horizon, October 2001*, 9(5)NCB University Press.
- [34] Reiser, R. A. (2001). A History of Instructional Design and Technology: Part I: A History of Instructional Design. *ETR&D*, 2001, Vol. 49, No. 1, 53–64.
- [35] Reimann, P. & Zumbach, J. (2003). Supporting virtual learning teams with dynamic feedback. In K. T. Lee & K. Mitchell (eds.), *The "Second Wave" of ICT in Education: from Facilitating Teaching and Learning to Engendering Education Reform* (pp. 424-430). Hong Kong: AACE.
- [36] Ridgway, J., McCusker, S., and Pead, D. (2004). Literature Review of e-Assessment. Bristol, UK: Nesta Future Lab.
- [37] Rottmann R. M. and Hudson H. T. (1983). Computer Grading As an Instructional Tool, *Journal of College Science Teaching*, Volume 12, Pages 152-156.
- [38] Tinoco, L., Fox, E., and Barnette, D. (1997). Online evaluation in WWW-based courseware. In *Proe. 2^{8th} SIGCSE Tech. Syrup.*, pp. 194-198.
- [39] WebCT Inc. Retrieved on 7 July 2008 from: www.webCT.com.
- [40] Woolley, D. R. (1994). PLATO: The emergence of on-line community. *Computer-Mediated Communication Magazine*, 1(3), 5.

Author(s):

Mohammad AL-Smadi Dr. techn student.
Graz University of Technology, Institute for Information Systems and Computer Media
Inffeldgasse 16c, A-8010 Graz
msmadi@iicm.edu

Christian Gütl Dipl.-Ing. Dr. techn
Graz University of Technology, Institute for Information Systems and Computer Media
Inffeldgasse 16c, A-8010 Graz
cguetl@iicm.edu