



## Expansion of Virtual E-Assessment Via Framework in Web-Based Courses

Marziyeh Nekoueizadeh, Taher Bahrani

Department of English Language, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran.  
(Corresponding author)

693mnekoeei@gmail.com

Department of English Language, Islamic Azad University, Mahshahr Branch, Mahshahr, Iran.  
taherbahrani@yahoo.com

### ABSTRACT

It is generally believed that assessment has important impact on instruction and learning (Gibbs, 1999; Scouler, 1998). The existence of gaps between educational system and the real world of work cause to materialize e-assessment as an essential component of teaching contexts. In fact, virtual and practical e-assessment provides feedbacks for students to meet what they need for regulation of good performance in the authentic environments. The expansion of virtual e-assessment is one of the controversial issues among training administrators to generate valid and reliable assessments results in higher education contexts by undertaking purposes such as providing a basis for evaluating the quality of instructional context and serving feedback loops for learners by associating themselves in self-assessment to actualize their desires. In the same line, the present paper has an attempt to discuss a conceptual framework for implementation of reliable and valid e-assessment in web-based courses by highlighting viewpoints of different scholars with regard to the role of e-assessment as an integral part of teaching and learning process.

### Indexing terms/Keywords:

E-assessment; Virtuality; Practicality; Framework; Educational system.



---

# Council for Innovative Research

Peer Review Research Publishing System

**Journal:** Journal of Advances in Linguistics

Vol 3, No. 1

[editor@cirworld.com](mailto:editor@cirworld.com)

[www.cirworld.com](http://www.cirworld.com), [member.cirworld.com](http://member.cirworld.com)



## DEFINITION OF E-ASSESSMENT

Traditionally, e-assessment uses computer and information technology to make the assessment process more efficient by automating functions that would otherwise require human assessors. These functions include item construction and selection, assessment delivery, item analysis, interpretation, scoring, and score presentation (Baker & O'Neil, 1995). E-Assessment is one of the domains of e-learning. It refers to assessment which is electronically delivered. The Joint Information Systems Committee (JISC, 2006) defines e-assessment as the end-to-end electronic assessment processes where Information and Communication Technology (ICT) is used for the presentation of assessment activity, and the recording of responses.

Writing in 1998 and looking to the future of e-assessment, Bennet (1999) describes three generations of e-assessment: First Generation tests using designs based closely on existing paper-based tests, Next-Generation tests which would use new formats including multimedia, constructed response, automatic item generation and automatic scoring, and Generation "R" (Reinvention) tests which would use complex simulations and intelligent tutors.

More recently, however, e-assessment refers to current efforts within the educational community to take advantage of technological advancements. These efforts are motivated by the need to go beyond "fill-in the bubble" tests and move forward to authentic assessments of complex skills and knowledge (Pellegrino, Chudowsky, & Glaser, 2001). Vendlinski and Stevens (2002) illustrate that technology provides new means to assess learning that will yield rich sources of data. E-assessment may include pre and post testing, diagnostic analysis, student tracking, rubric use/analysis, the support and delivery of authentic assessment through project based learning (e.g. WebQuests, simulations, e-portfolios), artifact collection, and data aggregation and analysis (Buzetto-More, 2006). The "e" in e-assessment involves making choices from a wide range of computer and information technologies that can be used as components of assessment development, implementation, and delivery. The word "assessment" in e-assessment relates to the purpose and features of the assessment. In general, the purpose of assessment can be categorized as formative, diagnostic, or summative. In addition, e-assessment must maintain the features of traditional assessment such as validity, reliability, fairness, and accessibility (Baker, O'Neil, & Linn, 1993; Shute, 2009).

## ASSOCIATED TERMS TO E-ASSESSMENT

Computer-based assessment (CBA), computer-assisted assessment (CAA), computer-aided assessment (CAA), computer adaptive assessment (CAA), computer adaptive testing (CAT), online assessment, technology-enabled assessment, technology-enhanced assessment, technology-embedded assessments, technology-based assessment, Internet-based testing, web-based assessment.

Lambert (2005) mentioned **some ways in which online technologies can be used in assessment including:**

- E-submission: assessments such as essays or reports are submitted electronically. Work in electronic format can be submitted to plagiarism detection services if required.
- E-Portfolios: collections of documents can be stored and delivered to viewers by a variety of means. They can often be commented on online. Specialist software is sometimes used for this, e.g. Pebble Pad
- Presentations: student presentations made using a variety of software applications, media and office tools can be shared and reviewed by single or multiple viewers.
- Peer Assessment: The ability to share materials electronically can be used to carry out projects which involve assessment by groups of peers.
- Self-Assessment: tools which aid reflection, audits of skills and online tests can be used to enable student to assess their progress.
- Electronic Voting Systems: used in lecture halls to enable immediate feedback onto the screen.

## REASONS FOR USING E-ASSESSMENT

Bull & McKenna, (2004) suggest a number of reasons that academics may wish to use Computer Assisted Assessment (CAA):

1. To increase the frequency of assessment, thereby:
  - \* motivating students to learn,
  - \* encouraging students to practice skills
2. To broaden the range of knowledge assessed
3. To increase feedback to students and lecturers
4. To extend the range of assessment methods
5. To increase objectivity and consistency
6. To decrease marking loads
7. To aid administrative efficiency



## QUESTIONS FORMS IN E-ASSESSMENT CONTEXT

E-assessment methods can provide different question/answer formats to support many types of test. E-assessment questions can be formatted in three ways.

### • **Closed-ended choice formats:**

- simple binary choices (yes/no, true/false);
- multiple-choice (choose one correct response);
- multiple-response (choose all correct responses);
- numeric entry (enter the correct number);
- slider (move pointer to the correct value); and
- hot spot (click on an image to identify a correct response) (Sangi, 2009).

### • **Open-ended formats:**

These usually require text-based answers varying from a single word or value to sentences and paragraphs; e.g.,

- fill in the blank (insert the missing words or values);
- short answer (free-form text field); and
- essay answer (long-response text field, likely to require manual grading) (Sangi, 2009).

### • **Ordering formats**

The question demands a response involving rank ordering: e.g.,

- drag and drop: place objects into the correct locations;
- order objects: rank objects according to the given criteria;
- match item: connect the objects in pairs; and
- connect the points: create an ordered connection of a set of points (Sangi, 2009).

## TEST FORMATS IN THE E-ASSESSMENT CONTEXT

### • **Attitude surveys**

Opinion-related assessment is obtained through attitudinal surveys. Attitudes in this context relate to different aspects of learning such as the course and its elements, the learning process, and the discipline. The simple Likert Scale or modifications of it are commonly used. Attitudinal surveys take many forms though usually involve a statement requiring a scaled agree/disagree response (Baggaley & Belawati, 2007).

### • **Student Assessment of Learning (SALG)**

It consists of statements about the degree of 'gain' (on a point scale) which students perceive they have made in aspects of the class. Instructors can add, delete, or edit questions. The instrument can be administered electronically with summary results made instantly available in statistical and graphic form. Such instruments are powerful tools and can easily be customized for efficient formative evaluation during a course or for faculty assessment (Sangi, & Malik, 2007).

### • **Conceptual diagnostic tests**

Students are often required to memorize theory, and can easily overlook sub-concepts. A conceptual diagnostic test presents items in a multiple-choice or short-answer format designed with common misconceptions in mind. Such tests can be applied to assess how well students understand and apply key concepts (Sangi, & Malik, 2007). Immediate feedback on the level of class understanding is available. Instructors have reported substantial improvements in class attendance and attitude toward the course through the use of such tests (Zeilik, 2007). Self-diagnostic tests can be created, and can aid in refining the thought process and overcoming biases. Diagnostic tests can easily be automated in an e-assessment mode, and can have a direct impact on the student's development (Sangi, & Malik, 2007).

### • **Performance assessment**

Performance assessments measure a student's ability to use specific knowledge and skills. They usually require the student to manipulate available knowledge to solve specific problems or to perform a conceptual analysis of various situations. Multiple-choice and fill-in-the-missing items questions can be used (Sangi, & Malik, 2007). Performance tests can be useful in disciplines involving scientific problems solving. However, a student's analytical growth can be difficult to measure accurately, and a more complete picture of student achievement can be achieved when performance assessments are used in conjunction with traditional forms of assessment (Sangi, & Malik, 2007).



### • **Rubrics**

Rubrics are tools for evaluating and providing guidance for students' writing. Rubrics can be translated to a digital format where they may be made available through an intranet or over the internet. Used for scoring, these scores provide meaningful assessment information. When connected to a database, they provide educators with data that can be aggregated (Buzzetto-More, 2006). Petkov and Petkova (2006) recommend course-embedded assessment as having the advantage of ease of implementation, low cost, timeliness, and student acceptance and note that the type of performance appraisal supported by rubrics is particularly effective when assessing problem solving, communication and team working skills. They explain that rubrics should not be considered checklists but rather criteria and rating scales for evaluation of a product or performance. They help to standardize assessment, provide useful data, and articulate goals and objectives to learners.

Petkov and Petkova (2006), who implemented rubrics in introductory IS courses, found that the use of rubrics helped to make assessment more uniform, better communicate expectations and performance to students, measure student progress over time, and help to lay the foundation for a long-term assessment program that combines projects and portfolios.

Andrade (2005) claimed that rubrics significantly enhance the learning process by providing both students and instructors with a clear understanding of the goals of the writing assignment and the scoring criteria. Stevens and Levi (2004) noted that rubrics facilitate timely and meaningful feedback to students. Peat (2006) suggested that, because of their explicitly defined criteria, rubrics lead to increased objectivity in the assessment of writing. Thus, different instructors might use a common rubric across courses and course sections to ensure consistent measurement of students' performance. Rubrics used in many subject areas in higher education generally include two elements: (a) a statement of criteria to be evaluated, and (b) an appropriate and relevant scoring system (Peat, 2006). Rubrics can be classified as either holistic or analytic (Moskal, 2000). Holistic rubrics award a single score based on the student's overall performance, whereas analytic rubrics give multiple scores along several dimensions. In analytic rubrics, the scores for each dimension can be summed for the final grade.

Rubrics (scoring tools) are a way of describing evaluation criteria (grading standards) based on the student's expected performance and outcomes. Typically, rubrics are used in grading written assignments or oral presentations, though they may be used to score any type of performance (Sangi, & Malik, 2007). Each rubric consists of a set of scoring criteria and point values associated with them. In most rubrics, the criteria are grouped into categories so that the instructor and student can identify the categories with specific levels of performance. In classroom use, the rubric aims to provide an objective external standard against which student performance can be compared. Rubrics sometimes generate conflicting scores on a given evaluation criterion, but generally provide an accessible means of communicating and developing assessment learning goals (Baggaley, 2009).

### • **Portfolios**

Portfolios are a collection of assignments done by the learner and evaluated by the instructor, to demonstrate mastery, comprehension, application, and synthesis of concepts. Portfolios can be used to assess learning-outcome achievement as well as to diagnose curriculum deficiencies that require improvement (Popper, 2005). Popper explained that portfolios should include a variety of samples of student work. According to Chun (2002), a portfolio should require students to collect, assemble, and reflect on samples that represent the culmination of their learning.

Cooper (1999) identified six considerations of the portfolio building process: identification of skill areas, design of measurable outcomes, identification of learning strategies, identification of performance indicators, collection of evidence, and assessment. Portfolio assessment strategies provide a useful structure for long-duration, in-depth assignments. The use of portfolio techniques transfers much of the responsibility for demonstrating concept mastery from the teacher to the student. Online methods of e-portfolio development are rapidly emerging (Sangi, & Malik, 2007). Portfolios are an effective form of alternative assessment that encourages students and educators to examine skills that may not be otherwise accessed using traditional means such as higher order thinking, communications, and collaborative abilities (Buzzetto-More, 2006; Wright, 2004).

A portfolio should require students to collect, assemble, and reflect on samples that represent the culmination of their learning (Chun, 2002) providing students with a diversity of opportunities to skills and abilities (Martell & Calderon, 2005). New High Tech High School in Napa, California involves students in long-term technology dependent portfolio projects (Page, 2006) and the business program at Merrimack College, has students engaged in an ongoing portfolio building process throughout their time in the program where students revisiting their portfolio will be able to observe their own growth (Popper, 2005).

Online portfolios have a number of advantages over those that are paper based as they support: a greater variety of artifacts and allow for greater learner expression; are dynamic and multimedia driven; accessible by a large audience; contain meta-documentation; easy to store; and may serve to promote a student academically or professionally (Buzzetto-More, 2006).

The collection of artifacts contained within a student learning portfolio can include text based items, graphic or multimedia driven elements, websites, and/or other items captured in an electronic format (Lorenzo & Ittelson, 2005a). According to Lorenzo and Ittelson (2005b) the skills required in the creation of electronic portfolios helps students learn, understand, and implement the information literacy process.



## AUTHENTIC ASSESSMENT

The earliest reference to authentic tests is likely that made by Archbald and Newman in 1988, in a book critical of standardized testing, that sought to promote assessment centered on meaningful real-world problems or tasks. Assessment is authentic when it measures products or performances that “have meaning or value beyond success in school” (Newman, Brandt & Wiggins, 1998). In fact, experts suggest that work being assessed should be authentic or based on the real world (Wiggins, 1990). Frequently known as project based learning, it is a form of instruction where students are immersed in challenging learning situations that are anchored in real world simulations (Page, 2006). According to Page, project based learning can support critical thinking, multilayered decision making, goal setting, problem solving, and collaboration. As a result, many institutions are anchoring their assessment activities into meaningful scenarios so that students are being assessed on their abilities to apply learning into realistic situations.

According to Newman et. al, (1998) assessments that ask questions and poses problems that have “real world” meaning to students meet one criterion for being authentic intellectual work, but there are two others related to disciplined inquiry that are unrelated to the realism of the assessment tasks. Wiggins was also an early proponent for the use of the term authentic to describe assessment with real-world application (1989). “Authentic refers to the situational or contextual realism of the proposed tasks” he has emphasized (Newman, Brandt & Wiggins, 1998). Terwilliger (1998) expressed concerns with Wiggins and others use of the term, viewing the label of authentic as a veiled criticism of traditional assessment approaches as somehow less authentic or inauthentic. Wiggins position is essentially that traditional assessment is not inauthentic, it is simply less direct and, probably, less meaningful to students (1989). Wiggins argues that traditional assessment is not faithful to the domains of performances and contexts that are most important for higher order thinking and learning (1993). As he used the term, authenticity is akin to fidelity.

In later work, Wiggins (1990, 1992, 1993) continued to emphasize that tasks should mirror real-world activities and elaborated that authentic tests should assess students’ “habits of mind”. Tasks are not authentic, necessarily, just because they are similar to real-world tasks, but they must mirror the complexity, collaboration, and high-level thinking that is necessary in the most intellectual of professional problem-solving and decision-making. The assessments act as instruction and skill-building opportunities, not merely as tools of evaluation.

To be authentic, the assessment must be part of a realistic learning context (Gulikers, Bastiaens & Martens, 2005) and make sense in the context of student work (Rennert-Ariev, 2005). “Authentic assessments are also contextualized; that is rather than assembling disconnected pieces of information, the tasks are set in a meaningful context that provides connections between real-world experiences and school-based ideas. These assessments are connected to students’ lives...”

The use of new technology in assessment issues related to pedagogy, content, and interaction. As these issues are addressed, there needs to be a subsequent alteration in the type of assessment used on such courses and the associated procedures (Waller, 2002).

According to Herrington and Herrington (1998) authentic assessment occurs within the context of an authentic activity with complex challenges, and centers on an active learner that produces refined results or products, and is associated with multiple learning indicators.

According to Olfos & Zulantay (2007).The authentic assessment not only evaluates the products, but also the processes involved. It is a process that monitors the learner’s progress using a variety of methods, such as observation records, interviews, and evidence gathering. It is consistent with Vygotsky’s dynamic assessment in that it is a mediated process in which social interaction stimulates the learning process. It also allows for collaborative learning.

Authentic assessment of educational achievement directly measures actual performance in the subject area. It was developed as a result of criticism of multiple-choice tests, which usually only provide a superficial idea of what a student has learned and do not indicate what a student can do with what was acquired (Aiken, 1996). One method to this end, according to Wiggins (1998), is the use of rubrics, which are sets of criteria that evaluate performance. Points are assigned according to how well each criterion is fulfilled, and are then used to provide the quantitative values.

## THE VALIDITY OF AUTHENTIC ASSESSMENT

Aiken (1996) plainly indicates the difficulty in establishing the validity and reliability of any authentic assessment. Schurr (1999) states that a disadvantage of authentic assessment is the difficulty, maybe even the impossibility, of obtaining consistency, objectivity, and/or standardization in its results. In the same vein, Linn, Baker and Dunbar (1991), hold that performance-based assessments are valid in terms of consequence, impartiality, transference, content coverage, cognitive complexity, significance, judgment, cost, and efficiency; consequently, reliability, understood as the stability of the results, is difficult to obtain.

Linn et al. state that “a greater dependence of critical judgments on the performance of tasks” is inevitable. This is a problem for large-scale assessments, but not for the smaller and more specific contexts to be found in superior education (Gipps, 1995), or in those cases in which the student is evaluated through class activities, where learning and evaluation are essentially one and the same (Reeves & Okey, 1996).

There are also problems with reliability in standardized tests. Gipps (1995) mentions national tests in the UK with reliability rates in posttests that are lower than those obtained in performance tests. Reeves and Okay (1996) state that authentic assessment takes place within a real-world context, where generalizations are of little value and, therefore, reproducibility



should not be a concern. Young (1995) adds that assessment needs can be perceived in a more functional way, and assessments can be validated in terms of their value in the real world rather than for their stability as instruments (Herrington and Herrington, 1998).

## **AUTHENTIC ASSESSMENT IN WEB BASED COURSES**

The two most important reasons for using authentic competency-based assessments are (a) their construct validity and (b) their impact on student learning, also called consequential validity (Gielen, Dochy, & Dierick, 2003). Construct validity of an assessment is related to whether an assessment measures what it is supposed to measure. With respect to competency assessment this means that (a) tasks must appropriately reflect the competency that needs to be assessed, (b) the content of an assessment involves authentic tasks that represent real-life problems of the knowledge domain assessed, and (c) the thinking processes that experts use to solve the problem in real life are also required by the assessment task (Gielen et al., 2003). Based on these criteria, authentic competency-based assessments have higher construct validity for measuring competencies than so-called objective or traditional tests have.

Consequential validity describes the intended and unintended effects of assessment on instruction or teaching (Biggs, 1996) and student learning (Dochy & McDowell, 1998). As stated, Biggs's (1996) theory of constructive alignment stresses that effective education requires instruction, learning, and assessment to be compatible.

If students perceive a mismatch between the messages of the instruction and the assessment, a positive impact on student learning is unlikely (Segers, Dierick, & Dochy, 2001). This impact of assessment on instruction and on student learning is corroborated by researchers as Frederiksen (1984, "The Real Test Bias"), Prodrômou (1995, "Backwash Effect"), Gibbs (1992, "Tail Wags the Dog"), and Sambell and McDowell (1998, "Hidden Curriculum"). Frederiksen (1984) and Prodrômou (1995) implied that tests have a strong influence on what is taught, because teachers teach to the test, even though the test might focus on things the teacher does not find most important.

Some see authentic assessment as a synonym for performance assessment (Hart, 1994; Torrance, 1995a), while others argue that authentic assessment puts a special emphasis on the realistic value of the task and the context (Herrington & Herrington, 1998). Reeves and Okey (1996) pointed out that the crucial difference between performance assessment and authentic assessment is the degree of fidelity of the task and the conditions under which the performance would normally occur. Authentic assessment focuses on high fidelity, whereas this is not as important an issue in performance assessment. These distinctions between performance and authentic assessment indicate that every authentic assessment is performance assessment, but not vice versa (Meyer, 1992).

Savery and Duffy (1995) defined authenticity of an assessment as the similarity between the cognitive demands—the thinking required—of the assessment and the cognitive demands in the criterion situation on which the assessment is based. A criterion situation reflects or simulates a real-life situation that could confront students in their internship or future professional life. Birenbaum (1996) further specified the competency concept by emphasizing that students need to develop not only cognitive competencies such as problem solving and critical thinking, but also meta-cognitive competencies such as reflection, and social competencies such as communication and collaboration.

In light of the constructive alignment theory (Biggs, 1996) authentic assessment should be aligned to authentic instruction in order to positively influence student learning.

Authentic assessment requires students to demonstrate relevant competencies through a significant, meaningful, and worthwhile accomplishment (Resnick, 1987; Wiggins, 1993).

Computing-mediated distance education introduces extraneous factors that could affect the validity of the course assessment system. One of these factors is identified as the usability of the web site. Usability deals with how well a system satisfies user needs and requirements. It applies to all the aspects of a system with which a user might interact, including installation and maintenance procedures (Orde, 2001). Another factor is attitude. Weller (2002) examines technical barriers in the assessment process of a web-based course, and points out the tension between individuality and robustness in submissions and the detection of plagiarism. Clarke et al. (2004) state that feedback to students on their assignments is an essential activity. They point out that tutorial support must be a crucial part of good distance course, where emails should be considered as a non-intrusive means of communication. Orde (2001) offers some suggestions to develop online courses, for instance: to consider a description of learners; to provide readily available technical support; to eliminate group activities easily done face-to-face; and to record and grade interactions such as e-mail and group discussion contributions. To Orde, the testing portion of Course Info requires that each quiz item be individually entered and submitted. If this feature of the software is used, the ID students advise allowing multiple attempts. To Orde, formative evaluation is an essential component and necessary to online course development.

Gatlin and Jacob (2002) discuss advantages of digital portfolios as part of one university's authentic pre-service teacher assessment. Fenwick and Parsons (1997) assert that effective assessment must be intricately woven throughout the teaching-learning process. Collaborative learning activities enable subjects to share their abilities and limitations, providing a better quality product than one that is the mere sum of individual contributions. Group interactions, following the indications for collaborative work given in a course, facilitate vicarious learning, which is hard for some subjects to experience if they do not interact with their peers (Bandura & Walters, 1963; Vygotsky, 1986).

Tools for objective testing, within virtual learning environments and within dedicated assessment engines (e.g. Question Mark Perception) allow teachers to orchestrate frequent assessment testing (e.g. online objective testing) which can be used both to offer flexibility in the time and place of assessment and/or to encourage students to spend more 'time on task'



out of class. Time on task has been correlated with enhanced student learning with many studies showing that frequent objective testing enhances the performance of students in final exams (e.g. Haigh, 2007).

Computer-supported assessment also makes it possible to enrich and make assessment tasks more authentic, for example, incorporating multimedia presentations and to enable better alignment of tests to student's levels of understanding (Conole & Warburton, 2005). For instance, as a test of their understanding, students learning a foreign language might watch a recorded video of current affairs program and answer objective questions relating to the content (Nicol, 2009). Adaptive testing involves modifying the nature of the test based on the responses the student has made to earlier tests. Although modifying interactions in this way can be achieved in paper tests this is far more efficient in computer-supported environments. Some researchers have also attempted to use computer programs in specific disciplinary domains (e.g. mathematics) to automatically generate multiple variations of the same class of problem types (Bennett, 1999).

## USE OF E-ASSESSMENT IN FORMATIVE ASSESSMENT

The nature and significance of formative e –assessment as an assessment has been studied from different standpoints. By considering that it is indicated as serial or repeated assessment and as practice assessment, it is relatively used as easy way for e-assessment within an educational system because it imposes less of a threat to educational system. Web-based formative assessment (WBFA) is WBA used by teachers for tracking students' progress to inform and adjust the educational process and to improve students' learning outcomes (Born, 2003). Hanson, Millington and Freewood (2001) and Mulligan (1999) noted that computerized tests, when used as objective tests with formative purposes, provide an efficient approach to learning since the use of frequent continuous feedback through such formative assessment has a positive impact on students' work.

However, many authors question the value of such an approach and argue for a view of formative assessment that places much more emphasis on the role of feedback in the learning process (e.g. Maughan, Peet, & Willmott 2001). Online formative assessment readily brings to mind the theory of Vygotsky (1986), where he suggests the need to build an atmosphere of social constructivism amongst the students and lecturers. The formative e-assessment process is supported through synchronous and asynchronous communication tools such as email, bulletin boards, news groups, and wikis. These electronic methods allow the students to pursue areas of perceived weakness and to affirm areas of strength (Challis, 2005).

In fact, formative e-assessment potentially has significant effect on learning occurring in higher education. Whitelock (2007), for example, argues for formative e-assessment as a means of promoting self-reflection and students taking control of their own learning, a view which echoes the ideas of Nichol (2006), who proposes a series of design principles for using formative e-assessment to support learner self-regulation. The core of a successful online formative assessment is the provision of immediate online feedback to students (Charman, 1999). Typically, this is a personalized feedback aimed at addressing the student's individual needs. In addition, the online feedback should provide explanations for wrong responses to questions or advices on approaching the questions differently.

The use of technology to support formative assessment has also been strongly indicated by Laurillard's Conversational Framework (Laurillard, 2002, 2007).

## USE OF E-ASSESSMENT IN SUMMATIVE ASSESSMENT

Online summative assessment is categorized as high-stake examinations which takes place at the end of a course of study. Online summative examinations enjoy the benefits of an internet-network environment, where student responses to questions are automatically uploaded and marked (Aojula et al, 2006). This online approach eliminates the need to safeguard printed questions and answers sheets. One of the key benefits of summative e-assessments is the ability to automatically mark tests, which can minimize human errors due to fatigue. Amidst the advantages, the summative e-assessment is perceived to be susceptible to a variety of security issues which challenges the online system. For example, to ensure data security, the item bank should be accessed by authorized parties only. Additionally, to ensure user security, the assessment should be delivered and taken by the correct student only. Hence, due to the high-stake nature of the summative tests it is essential to maintain a high security level.

## The CHARACTERISTIC OF AUTHENTIC ASSESSMENT

Many authors have provided criteria with which to design and evaluate authentic assessment. For example, Newmann and Wehlage (1993), Wiggins (1990, 1993), Reeves (2000), Reeves and Okey (1996) and others have provided guidelines or elements that help to explain the nature of authentic assessment. Building on our summary of the essential characteristics of authentic assessment (Herrington & Herrington, 1998), the below list attempts to provide a synthesis of the recent literature and research while considering assessment context, student factors, task factors and indicators. Using these guidelines, assessment is most likely to be authentic if it satisfies the following criteria:

### **Context:**

- Requires fidelity of the task to the conditions under which the performance would normally occur (Reeves & Okey, 1996; Meyer, 1992; Wiggins, 1993)
- Requires connectedness and transfer to the world beyond the classroom (Newmann & Wehlage, 1993; Newmann & Archbald, 1992)

**Student factors:**

- Requires problem solving skills and higher order thinking (Reeves, 2000; Newmann & Wehlage, 1993)
- Requires production of knowledge rather than reproduction (Newmann & Archbald, 1992) •Requires significant student time and effort in collaboration with others (Linn, Baker, & Dunbar, 1991) (Reeves, 2000)
- Is characterized by substantive conversation (Newmann & Wehlage, 1993)
- Requires students to be effective performers with acquired knowledge, and to craft polished, performances or products (Wiggins, 1990, 1993)
- Promotes depth of knowledge (Newmann & Wehlage, 1993)

**Task factors:**

- Stimulates a wide range of active responses (Reeves, 2000)
- Involves complex, ill structured challenges that require judgment, multiple steps, and a full array of tasks (Wiggins, 1990, 1993; Linn, Baker, & Dunbar, 1991; Torrance, 1995b) (Reeves, 2000)
- Requires the assessment to be seamlessly integrated with the activity (Reeves & Okey, 1996; Young, 1995).

**Indicators**

- Provides multiple indicators of learning (Lajoie, 1991; Linn, Baker, & Dunbar, 1991)
- Achieves validity and reliability with appropriate criteria for scoring varied products (Wiggins, 1990; Lajoie, 1991; Resnick & Resnick, 1992)

Such guidelines enable teachers to create learning environments using authentic contexts and scenarios that ensure assessment truly measures whether students can use their knowledge effectively and realistically, as opposed to the reproduction of surface knowledge that is quickly forgotten after an examination or test.

**DISADVANTAGE USE OF AUTHENTIC ASSESSMENT**

Some critics claim authentic assessment fails to remedy the problems of standardized testing. Authentic assessments can be as contrived as their traditional counterparts (Terwilliger, 1996, p.6). "There is no absolute distinction between performance tests and other classes of tests" (Terwilliger, 1996, p.13).

Promotion of authentic assessment is primarily based on the failures of standardized tests. The academic world, however, does not unanimously denounce traditional testing. "While traditional exams are imperfect, these tests effectively provide information about student standing relative to peers on the domains tested" (Archbald, 1991). Some critics believe authentic tests cannot require the same factual knowledge as traditional assessment (Terwilliger, 1996). Authentic assessment may claim to address some problems in the current classroom test, but not without presenting its own dilemmas.

**A FRAMEWORK FOR THE DEVELOPMENT OF ASSESSMENT IN AN ONLINE CONTEXT**

Basically, language tests need to meet three general principles that underlie all good test design: they need to be valid, reliable, and practical (Weir, 1993). In a more detailed model for establishing the usefulness of language tests, Bachman & Palmer (1996; 2010) consider the following six qualities: (1) reliability , or the consistency of measurement; (2) construct validity , or the degree to which scores can be interpreted as indicators of the measured construct, for instance oral proficiency in the foreign language; (3) authenticity , or the extent to which a given test task reflects the characteristics of real-world tasks; (4) interactiveness, or the extent to which a given test task engages the learner's competences and strategies; (5) impact , or the effects of the test on individuals, teaching and learning activities, educational systems, and society in general; and (6) practicality , or the question whether design, development, and use require more resources than are available.

Computer-supported assessment also makes it possible to enrich and make assessment tasks more authentic, for example, incorporating multimedia presentations and to enable better alignment of tests to student's levels of understanding (Conole and Warburton, 2005).

The below presented dimensions and criteria, can be used as a framework for the development of a competence assessment program in an online context. The suggested dimensions and criteria by Tinoca (2012) cover not only the six aspects of construct validity proposed by Messick (1994, 1995), but also the ten quality criteria for competence assessment programs suggested by Baartman et al. (2007).

In fact, the discussed criteria, more than just illustrating the different features of each dimension, allow for the operational description of each criterion stage of implementation, and so contribute to the evaluation of the achieved assessment strategy quality level.



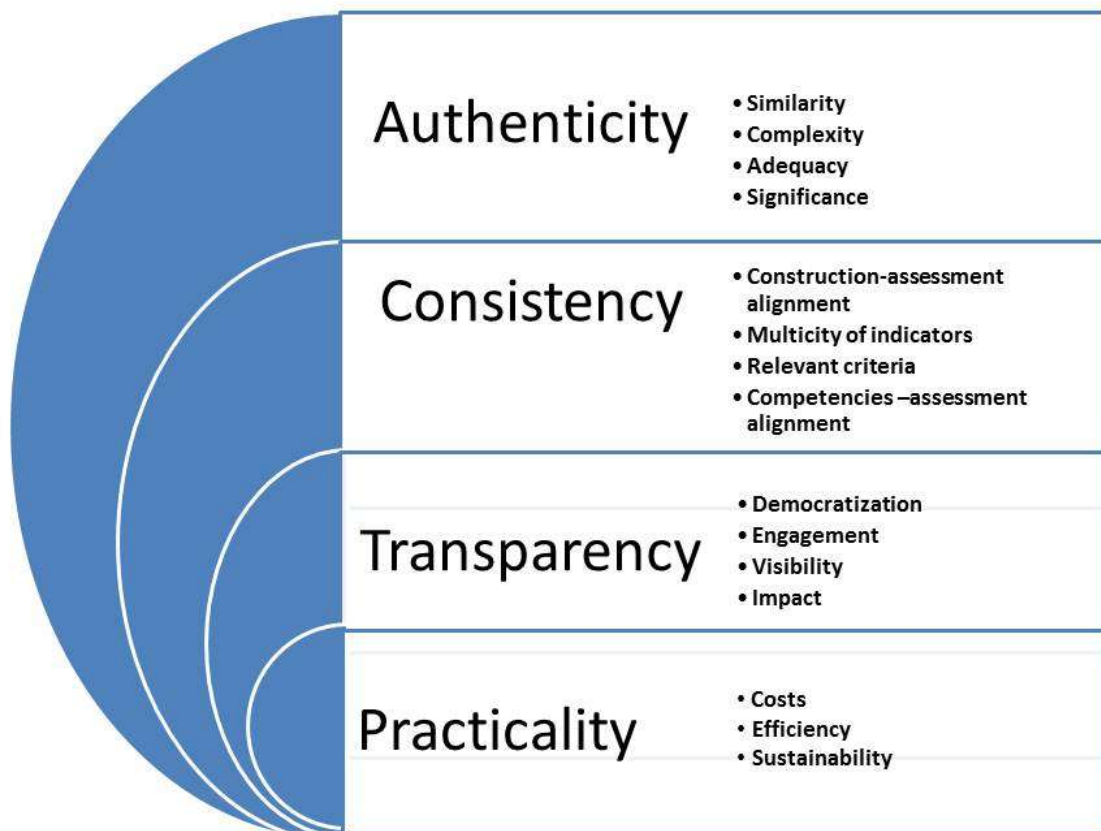


Fig 1: .A conceptual framework for e-assessment suggested by Tinoca (2012).

The concept of **AUTHENTICITY** is related to the degree of similarity between the competences being assessed and the ones required in real/professional life. Wiggins (1990) suggests that work being assessed should be authentic or based on the real world.

There are four reference criteria contributing to the degree of assessment authenticity:

- **Similarity** – refers to the way in which the assessment strategy is related to the real life context (physical and social), meaning that assessment should reflect the competences needed in real/professional life (Baartman et al., 2007; Dierick and Dochy, 2001; Gulikers et al., 2004). The physical context refers to the type and number of available digital resources, while the social context is supposed to be aligned with the equivalent social processes in a real/professional situation.
- **Complexity** – refers to the nature of the e-assessment tasks, more specifically, to the cognitive challenges that are imposed by its resolution/development, similar to real/professional, daily life challenges often ill -structured and with a variety of possible solutions (Gulikers et al., 2004; Herrington & Herrington, 1998).
- **Adequacy** – is connected with the need to provide adequate performing conditions (time, digital resources, etc.) for the e-assessment tasks, in accordance with their complexity. This often implies the removal of unrealistic restraints imposed by formal educational contexts (Gulikers et al., 2004; Herrington & Herrington, 1998). The participants' responses should explain understandings and conclusions; they should be clear, coherent and provide richness in details, qualifications and argument (MacLellan, 2004).
- **Significance** – includes the significant value of the e -assessment task for students, instructors and employers (Baartman et al., 2007; Gulikers et al., 2004). The connection between the e -assessment tasks and the learning needs should be clear and perceived by the students/learners.

The **CONSISTENCY** dimension is considered as an answer to the traditional demands for validity and reliability, associated with psychometric indicators. It takes into account that the assessment of competences requires the implication of a variety of assessment methods, in diverse contexts, by different assessors, as well as the adequacy of the employed strategies (Dierick & Dochy, 2001).

This dimension is comprised by four criteria:



- **Construction -assessment alignment** – refers to the need to provide e-assessment scenarios that are representative of the learning situations experienced by the students (Palm, 2008), warranting in this way the agreement between the work developed during the instruction and the assessment tasks being used.
- **Multiplicity of indicators** – this criterion is related to the need of employing a variety of e-assessment methods, contexts, moments, and assessors (Dierick & Dochy, 2001; Herrington & Herrington, 1998) during learning process.
- **Relevant criteria** – considers the relevance of the assessment criteria being used for the competences that are being assessed (individual or collaborative (Herrington & Herrington, 1998; Pereira, Tinoca, & Oliveira, 2010).
- **Competences-assessment alignment**– is related to the need of assuring the coherency between the competences that are intended to be developed and the assessment design being used (Palm, 2008; Pereira, Tinoca, & Oliveira, 2010).

The **TRANSPARENCY** dimension intends to make the entire competence assessment program visible and comprehensible for all participants. For this reason, it is important that the students/learners are able to understand the fairness of their assessment, as precisely as their instructors, requiring them to have a complete knowledge of all assessment criteria and their relative weights (Tinoca, 2012). According to Dierick and Dochy (2001), the clarification of the assessment criteria coupled with the fact that the students/learners have prior knowledge of the competence assessment program has a positive impact on their learning. For this dimension four criteria were considered:

- **Democratization**– refers to the availability and possible participation in the definition of the assessment criteria (Dierick & Dochy, 2001). Moreover, the students/learners should know, from the onset, what the assessment goals are and who their assessors are going to be. This way, they know immediately what is expected from them, and can adjust their learning processes accordingly (McConnell, 2006).
- **Engagement**– is related to the availability and possible participation in the definition of the learning goals (MacLellan, 2004b) and performance criteria (Pereira, Tinoca, & Oliveira, 2010). It allows students/learners to participate in the definition of their learning environment, further contributing to their active participation, commitment and responsibility (Pereira, Tinoca, & Oliveira, 2010).
- **Visibility**– refers to the possibility of presenting/sharing their learning processes and/or products with others (peers, assessors, instructors, community, etc.) (Gulikers et al., 2004).
- **Impact** – is related to the effects that the e-assessment strategies have in the learning process and in the design of the educational program (Baartman et al., 2007). According to Brinke (2008) the assessment design must have a positive impact in the learning process.

The **PRACTICALITY** dimension is related with the feasibility of the assessment strategy. This dimension is particularly important when designing a competence assessment program given the complexity of its design. It implies an effective management in terms of time and cost/efficiency balance for both assessors and organizations (Brown, 2004). Furthermore, it should assure that the students /learners consider the assessment tasks as being doable, relevant and contributing to their learning. It is composed of three criteria:

- **Costs**– referring to the time costs (for both assessors and students/learners) (Brinke, 2008; Dierick & Dochy, 2001), as well as the digital resources or additional investments, particularly training, that may be necessary to implement the assessment design.
- **Efficiency** – goes beyond costs, to consider the relationship costs/effects of the e-assessment design for the institutions, instructors and learners, particularly considering the expected results. It should, for example, be translated into the choice of certain technological tools aimed at promoting a more efficient assessment (Brown, 2004).
- **Sustainability** – is related to the need of assuring that it is possible to successfully implement and sustain the proposed assessment design, taking into account the learner profiles (level of education, previous training, familiarity with the assessment tools, prior knowledge, competences, etc.) and the contextual constraints, both for the organizations and for the assessors.

## CONCLUSION

It can be concluded that, although the importance of e-assessment has been emphasized by the majority of the researchers, little has also been written about how to implement virtual and practical e-assessment in educational contexts for preparing learners when encountered to real-life situations. In nutshell, designing basic framework for ideally implementation and administration of e-assessment system in any platform should be carefully considered from different aspects in order to eliminate any potential pitfalls in application of e-assessment as new technology.

## REFERENCES

- [1] Aiken, L. R. (1996). *Rating scales and checklists*. Canada: John Wiley sons, Inc.
- [2] Aojula, H., Barber J., Cullen R., & Andrews J. (2006). Computer-based online summative assessment in pharmacy teaching. *pharmacy education: The Manchester experience*. *Pharmacy Education*, 6(4), 229-236. Sannella, M. J. 1994



Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

- [3] Archbald, D. A. (1991). *Authentic assessment: What it means and how it can help schools*. Madison, WI: national center for effective schools research and development (ERIC Document Reproduction Service No. ED 366 624). Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.
- [4] Archbald, D. A., & Newmann, F. M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. Washington, DC. : Office of Educational Research and Improvement. Spector, A. Z. 1989. *Achieving application requirements*. In *Distributed Systems*, S. Mullender
- [5] Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Vleuten, C. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, 2, 114-129.
- [6] Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- [7] Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University.
- [8] Baggaley J. & Belawati, T. (Eds.) (2007). *Distance Education Technology in Asia. Part 1: Past and present*. Lahore: Virtual University of Pakistan.
- [9] Baggaley J. (Ed.) (2009). *Distance Education Technology in Asia*. Lahore: Virtual University of Pakistan.
- [10] Baker, E. L., & O'Neil, H. F., Jr. (1995). Computer technology futures for the improvement of assessment. *Journal of Science Education and Technology*, 4(1), 37-45.
- [11] Baker, E. L., O'Neil, H. F., Jr. & Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48(12), 1210-1218.
- [12] Bandura, A., & Walters, R. H. (1963). *Social Learning and Personality Development*. New York: Holt, Rinehart and Winston.
- [13] Bennett, R. E. (1999). Using new technology to improve assessment. *Educational Measurement: Issues and Practice*, 18(3), 5-12.
- [14] Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.
- [15] Birenbaum, M. (1996). *Assessment 2000: Towards a pluralistic approach to assessment*. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 3-29). Boston, MA: Kluwer Academic Publishers.
- [16] Born, A. D. (2003). *Web-based student assessment*. In A. K. Aggarwal (Ed.), *Web-based education: Learning from experience* (pp. 165-188). Hershey, USA: Idea Group Inc.
- [17] Brinke, D. (2008). *Assessment of prior learning*. Maastricht, the Netherlands: Datawyse.
- [18] Brown, S. (2004). *Assessment for learning*. *Learning and Teaching in Higher Education*, 1, 81-89.
- [19] Buzzetto-More, N. A. (2006). *The e-Learning and business education paradigm: Enhancing education, assessment, and accountability*. *Proceedings of the Maryland Business Education Association Conference*. Ocean City: MD.
- [20] Challis, D. (2005). Committing to quality learning through adaptive online assessment. *Assessment in Education*, 30(5), 519-527.
- [21] Charman, D. (1999). *Issues and impacts of using computer-based assessments (CBAs) for formative assessment*. In S. Brown, P. Race & J. Bull (Eds.), *Computer assisted assessment in higher education* (pp. 85-93). Abingdon, Oxon: Routledge.
- [22] Chun, M. (2002). Looking where the light is better: A review of the literature on assessing higher education quality. *Peer Review* 4(2-3), 16-25.
- [23] Clarke, M., Butler, C., & Schmidt-Hansen, P. (2004). Quality assurance for distance learning: A case study at Brunel University. *British Journal of Educational Technology*, 35(1), 5-11.
- [24] Conole, G., & Warburton, B. (2005). A review of computer-assisted assessment. *ALT-J Research in Learning Technology*, 13(1), 17-31.
- [25] Cooper, T. (1999). *Portfolio assessment: A guide for lecturers, teachers, and course designers*. Perth, Australia: Praxis Education.
- [26] Dierick, S., & Dochy, F. J. R. C. (2001). New lines in edumetrics: New forms of assessment lead to new assessment criteria. *Studies in Educational Evaluation*, 27, 307-329.
- [27] Dochy, F. J. R. C., & McDowell, L. (1998). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279-298.



- [28] Fenwick, T., & Parsons, J. (1998). Starting with our stories: Towards more authentic assessment in adult education. *Adult Learning*, 26(4), 25-30.
- [29] Frederiksen, N. (1984). The real test bias, influences of testing and teaching on learning. *American Psychologist*, 39(3), 193-202.
- [30] Gatlin, L., & Jacob, S. (2002). Standards-Based digital portfolios: A component of authentic assessment for pre-service teachers. *Action in Teacher Education*, 23(4), 28-34.
- [31] Gibbs, G. (1999). Using Assessment Strategically to Change the Way Students Learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: choosing and using diverse approaches* (pp.41-53). Buckingham, UK: SRHE and Open University Press.
- [32] Gielen, S., Dochy, F., & Dierick, S. (2003). Evaluating the consequential validity of new modes of assessment: The influence of assessment on learning, including the pre-, post-, and true assessment effects. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimizing new modes of assessment: In search of quality and standards* (pp. 37-54). Dordrecht: Kluwer Academic Publishers.
- [33] Gipps, C. (1995). What do we mean by equity in relation to assessment? *Assessment in Education*, 2(3), 271-281.
- [34] Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Design*, 53, 67-87.
- [35] Gulikers, J., Bastiaens, T., & Martens, R. (2005). The surplus value of an authentic learning environment. *Computers in Human Behavior*, 21, 509-521.
- [36] Haigh, M (2007). Sustaining learning through assessment: An evaluation of the value of a weekly class quiz. *Assessment and Evaluation in Higher Education*, 32(4), 457-474.
- [37] Hanson, J., Millington, C. & Freewood, M. (2001). Developing a methodology for online feedback and assessment. In M. Danson (Ed.) *Fifth International Computer Assisted Assessment (CAA) Conference*. Leicestershire, UK, Loughborough University.
- [38] Hart, D. (1994). *Authentic assessment: A handbook for education*. Menlo Park, CA: Addison-Wesley Publishing Company.
- [39] Herrington, J. & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research & Development*, 17(3), 305-322.
- [40] JISC (2006). *Roadmap for e-assessment*. Milton Keynes, UK: Open University Press.
- [41] Lajoie, S. (1991). A framework for authentic assessment in mathematics. *NCRMSE Research Review: The Teaching and Learning of Mathematics*, 1(1), 6-12.
- [42] Laurillard, D. (2002). *Rethinking university teaching: A conversational framework for the effective use of learning technologies* (2nd ed.). London, UK: Routledge.
- [43] Laurillard, D. (2007). Pedagogical forms for mobile learning: Framing research Question. In N. Pachler (Ed.), *Mobile learning: Towards a research Agenda* (pp. 153-175). London: WLE Centre, IoE.
- [44] Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment. *Educational Researcher*, 20(8), 15-21.
- [45] Lorenzo, G, & Ittelson, J. (2005a). *An overview of e-portfolios*. Boulder, CO: Educause Learning Initiative.
- [46] Lorenzo, G, & Ittelson, J. (2005b). *Demonstrating and assessing student learning with e-portfolios*. Boulder, CO: EduCause Learning Initiative.
- [47] Maclellan, E. (2004). How convincing is alternative assessment for use in higher education? *Assessment & Evaluation in Higher Education*, 29(3), 311-321.
- [48] 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 (Eds.), *Assessment of student learning in business schools: Best practices each step of the way* (pp. 1-22). Tallahassee, Florida: Association for Institutional Research.
- [49] Maughan, S., Peet, D., & Willmott, A. (2001). On-line formative assessment item banking and learning support. In M. Danson & C. Eabry (Eds.) *5th International Computer Assisted Assessment (CAA) Conference*. Loughborough, UK: Loughborough University.
- [50] McConnell, D. (2006). *E -learning groups and communities*. Berkshire, UK: Open University.
- [51] Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13-23.
- [52] Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.



- [53] Meyer, C. (1992). What's the difference between authentic and performance assessment? *Educational Leadership*, 49(8), 39-40.
- [54] Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment. Research & Evaluation*, 7(3), 1-7.
- [55] Mulligan, B. (1999). Pilot study on the impact of frequent computerized assessment on student work rates, In M. Danson, & R. Sherratt (Eds.) *Proceedings of the 3rd Annual Computer Assisted Assessment (CAA) Conference*. Loughborough, UK: Loughborough University.
- [56] Newmann, F. M., & Archbald, D. A. (1992). The nature of authentic academic achievement. In H. Berlak, F. M. Newmann, E. Adams, D. A. Archbald, T. Burgess, J. Raven & T. A. Romberg (Eds.), *Toward a new science of educational testing and assessment* (pp. 71-84). Albany, NY: State University of New York Press.
- [57] Newmann, F. M., & Wehlage, G. (1993). Five standards of authentic instruction. *Educational Leadership*, 50(7), 8-12.
- [58] Newmann, F. M., Brandt, R., & Wiggins, G. (1998). An exchange of views on semantics, psychometrics, and assessment reform: A close look at 'authentic' assessments. *Educational Researcher*, 27(6), 19-22.
- [59] Nicol, D. J. (2009). Assessment for learner self-regulation: Enhancing achievement in the first year using learning technologies. *Assessment and Evaluation in Higher Education*, 34(3), 335-352.
- [60] Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- [61] Olfos, R., & Zulantay, H. (2007). Reliability and validity of authentic assessment in a web based course. *Educational Technology & Society*, 10(4), 156-173.
- [62] Orde, B. (2001). Online course development: Summative reflections. *International Journal of Instructional Media*, 28(4), 397-403.
- [63] Page, D. (2006). 25 tools, technologies, and best practices. *T. H. E. Journal*, 33(8), 42-46.
- [64] Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment Research & Evaluation*, 13(4), 1-11.
- [65] Peat, B. (2006). Integrating writing and research skills: Development and testing of a rubric to measure student outcomes. *Journal of Public Affairs Education*, 12, 295-311.
- [66] Pellegrino, J., Chudowsky, N., & Glaser, R. (Eds.) (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- [67] Pereira, A., Tinoca, L., & Oliveira, I. (2010). Authentic assessment contribution to competence based education at Universidade Aberta: Questions and challenges. In S. Mukerji & P. Tripathi (Eds.), *Cases on technological adaptability and transnational learning: Issues and challenges*: IGI Global.
- [68] Petkov, D., & Petkova, O. (2006). Development of scoring rubrics for IS projects as an assessment tool. *Issues in Informing Science and Information Technology Education*, 3, 499-510.
- [69] Popper, E. (2005). Learning goals: The foundation of curriculum development and assessment. In K. Martell & T. Calderon (Eds.), *Assessment of student learning in business schools: Best practices each step of the way* (pp. 1-23). Tallahassee, Florida: Association for Institutional Research.
- [70] Prodromou, L. (1995). The backwash effect: From testing to teaching. *ELT Journal*, 49(1), 13-25.
- [71] Reeves, T. C. (2000). Alternative assessment approaches for online learning environments in higher education. *Journal of Educational Computing Research*, 23(1), 101-111.
- [72] Reeves, T. C., & Okey, J. R. (1996). Alternative assessment for constructivist learning environments. In B. G. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design* (pp. 191-202). Englewood Cliffs, NJ: Educational Technology Publications.
- [73] Rennert-Ariev, P., (2005). A theoretical model for the authentic assessment of teaching: Practical assessment, research and evaluation. *A Peer-reviewed electronic journal*, 10(2), 1-11.
- [74] Resnick, L. B. (1987). Learning in school and out. *Educational Leadership*, 16(9), 13-20.
- [75] Resnick, L. B., & Resnick, D.P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessment: Alternative views of aptitude, achievement and instruction* (pp. 37-75). Boston: Kluwer.
- [76] Sangi, N. A., & Malik, I. (2007). E-assessment models and methods for student evaluation. In Editor (first initial, last name) (Ed.) *21st Asian Association of Open Universities Conference*. Putrajaya:Malaysia:Kuala Lumpur (Malaysia) university.
- [77] Savery, J., & Duffy, T. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology*, 35, 31-38.



- [78] Schurr, S. (1999). *Authentic Assessment from A to Z*. USA: National Middle School Association.
- [79] Scouller, k. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35, 453-427.
- [80] Segers, M., Dierick, S., & Dochy, F. (2001). Quality standards for new modes of assessment: An exploratory study of the consequential validity of the Overall test. *European Journal of Psychology of Education*, 16(4), 569-586.
- [81] Shute, V. J. (2009). Simply assessment. *International Journal of Learning, and Media*, 1(2), 1-11.
- [82] Stevens, D. D., & Levi, A. (2004). *Introduction to rubrics: An assessment tool to save grading time, convey feedback, and promote student learning*. Sterling, VA: Stylus.
- [83] Terwilliger, J. S. (1996). *Semantics, psychometrics, and assessment reform: A close look at 'Authentic' tests*. New York, NY: National Council on Measurement in Education.
- [84] Terwilliger, J. S. (1998). Rejoinder: Response to Wiggins and Newman. *Educational Researcher*, 27(6), 22-23.
- [85] Tinoca, L. (2012). Promoting e-assessment quality in higher education: A case study in online professional development. In K. Fernstrom (Ed.) *Proceedings of International conference on information communication technologies in education* (pp. 213- 223). Portugal: University of Lisbon.
- [86] Torrance, H. (1995a). Introduction. In H. Torrance (Ed.), *Evaluating authentic assessment: Problems and possibilities in new approaches to assessment* (pp. 1-8). Buckingham, UK: Open University Press.
- [87] Torrance, H. (1995b). *Evaluating authentic assessment: Problems and possibilities in new approaches to assessment*. Buckingham, UK: Open University Press.
- [88] Vendlinski, T., & Stevens, R. (2002). Assessing student problem-solving skills with complex computer based tasks. *Journal of Technology, Learning, and Assessment*, 1(3), 1-21.
- [89] Vygotsky, L. S. (1986). *Thought and Language*. Cambridge, MA: MIT Press.
- [90] Weir, C. (1993). *Understanding & developing language tests*. Hemel Hempstead, UK: Prentice Hall.
- [91] Weller M. (2002). Assessment issues on a web-based course. *Assessment & Evaluation in Higher Education*, 27(2), 109-116.
- [92] Whitelock, D. (2007). Computer assisted formative assessment: Supporting students to become more reflective learners. In C. P. Constantinou, Z. C. Zacharia & M. Papaevripidou (Eds.) *8th International Conference on Computer Based Learning in Science* (pp. 492-504). Crete, Greece: E-Media, University of Crete.
- [93] Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70(9), 703-716.
- [94] Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, 2(2).
- [95] Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership*, 49(8), 26-34.
- [96] Wiggins, G. (1993). Assessment to improve performance, not just monitor it: Assessment reform in the social sciences. *Social Science Record*, 30(2), 5-12.
- [97] Wiggins, G. (1996). Practicing what we preach in designing authentic assessments. *Educational Leadership*, 54(4), 18-25.
- [98] Wright, B. (2004). *An assessment planning primer: Getting started at your own institution*. 13th Annual Northeast Regional Teaching Workshop.
- [99] Young, M.F. (1995). Assessment of situated learning using computer environments. *Journal of Science Education and Technology*, 4(1), 89-96.
- [100] Zeilik, M. (2007). *Conceptual Diagnostic Tests*. University of New Mexico: Department of Physics and Astronomy.