

530945-LLP-1-2012-1-GR-KA3-KA3MP

CRITON - Prediction of e-learners' progress and timely assessment of the achievement of learning outcomes in Lifelong Learning



D2.1 Study of Assessment practice in e-learning environments

Project:	CRITON - Prediction of e-learners' progress and timely assessment of the achievement of learning outcomes in Lifelong Learning
Work package:	WP2 Comparative analysis of assessment methods in eLearning
Lead Participant:	HOU, Greece
Authors:	Adamantia Spanaka, Achileas Kameas
Document Type:	Document
Distribution:	Project consortium, EACEA
Status:	Final
Document file:	WP2_1_FINAL_DELIVERABLE.docx
Version:	V2
Date:	18.12.2013
Number of pages:	68

INDEX

1	Introduction: Assessment in e-learning environments around Europe.....	4
2	From Assessment to e-Assessment.....	6
2.1	Some important considerations of e-assessment.....	7

PART I: THEORETICAL FRAMEWORK

3	Assessment: Conceptual Approach.....	9
4	Diagnostic, Formative and Summative assessment.....	10
5	Low-Value Assessments.....	11
6	High-Value Assessments.....	12
7	Authentic Assessment.....	12
8	Why there is a Need for Assessment Questions?.....	13
9	Question Types in Assessments.....	13
9.1	Multiple Choice.....	15
9.2	True/False.....	17
9.3	Essay.....	17
9.4	Matching/ Reordering/Rearrangement.....	18
9.5	Completion.....	18
9.6	Presentation/Portfolio.....	19
9.7	Question types in Assessments: Conclusions.....	21
10	Different (or Impossible) Online Assessments and Possible Mitigation.....	22
11	Methods for Formulating the Right Mix of Assessments.....	22
12	Feedback in formative assessment.....	23

13	SOME MORE BASIC DIDACTICAL ASPECTS OF ASSESMENT ACTIVITIES	25
13.1	Learning objectives	25
13.2	Criteria for success	26

PART II

LITERATURE REVIEW

14	The case of Germany: Lessons Learned - 7 years of e-assessment	27
15	The case of Austria: e-Assessment practices	29
16	The case of Greece: e-Assessment literature review	34
17	The case of Lithuania: Assessment literature review in conventional education and in eLearning environments	40
18	The case of Finland: A Success Story	46
19	A global view: Application of formative assessment to eLearning	50
20	Pedagogic approaches of e-Learning: The big gap in the literature review	51
21	Benchmarking models	51
22	Final Conclusions: Restrictions and Implementation of Assessment in the E-Learning Situation	51

PART III

FINAL REMARKS

24	A Conceptual Classification of Assessment	54
	References	58
	ANNEX	65
	Glossary	66

1 Introduction: Assessment in eLearning environments around Europe

*E-Learning*¹ can be considered to be highly related to learning and teaching (Gruttmann & Usenev, 2011; Raith, 2009; Jain et al., 2002). Therefore, pedagogy and didactic are important aspects for all facets of eLearning (Spector & Ohrazda, 2003), expanding from the creation of the courseware and application of an eLearning system to the evaluation of the learning progress. In particular, didactic has a great impact on designing and mediating learning content (e.g. see Gunawardena & Mclsaac, 2003). Considering the traditional teaching process, the teacher has to decide which competencies have to be mastered by the students, to which point they extend and how the result of the learning process should be measured. Such considerations are usually realised by determining learning objectives and assessment methods for a course.

With respect to Mödritscher & Sindler (2005), applying certain didactical principles can be challenging in the eLearning situation. In particular, the *assessment*² takes an important place during the development of any of e-learning course. It includes a variety of activities such as testing, problem solving, developing collaborative or individual projects development, participating in discussions etc. The object of assessment in the educational process is not only the knowledge. As IDS (2002) points out, the necessity of using the appropriate assessment which should be performed not only to grade students, but also to measure the learning process.

The assessment area is enlarged toward the skills and competences. According to Mödritscher et al. (2004) the competence is a mastering of knowledge and skills at level that is sufficient for their application for doing concrete work. Howard Gardner's Multiple Intelligences Theory (1983) describes three main classes of competencies:

- (1) knowledge seen as a kind of mental model about parts of the real world,
- (2) skills related to the capacity of applying and using acquired knowledge, and
- (3) attitudes dealing with social or affective aspects.

In praxis, competency is supposed to be related to more than one of these classes. In most

^{1,2} The glossary at the end of the report defines this term.

cases a strong focus on one class can be recognised, but adequately mixed competencies can also be found. That's why the assessment method has a great impact on the students' learning behavior (Rust, 2002). And this is the reason, the eLearning assessor should keep in mind that when eLearning is implemented, students should be led to the desirable direction as they are taught in the classroom. Both classroom teaching and eLearning should be comparable and yield significant impacts to students' progress.

So, the approach of using different assessment methods should be applied in eLearning environments. The *assessment methods*³ are based on relevant assessment activities. These activities could be aggregated in an assessment unit and they could be used to measure student's achievements and driving of learning path in a concrete e-learning course, module, chapter, lesson etc. Teachers have to consider which type of question they use for assessment depending on the level of learning objectives, size of the class, reliability in grading, prevention of cheating, exam construction and grading time, and several other criteria.

Thus, three critical questions arise when planning and evaluating a course: What competencies should be mediated to the students? To which extend should these competencies be mastered? And how can the results of the learning process be measured after having finished the course?

Based on all these, different assessment methods used in eLearning environments around Europe are in the focus of this study.

There will also be an answer to questions like:

1. What are the basic types of assessments available and which are the most widely used in Europe?
2. What is the difference between low-value and high-value assessments? Formative and summative assessments? When each one should be used in eLearning environment?
3. What are some different methods of assessment in eLearning environment?
4. Are there assessment types which are difficult or currently "impossible" to be created in

³ The glossary at the end of the report defines this term.

eLearning environment?

5. What are some methods to find the right mix of assessments for an online course or learning sequence?

6. How can feedback influence student achievement in eLearning?

2 From Assessment to e-Assessment

Technology can add value to assessment practice in a multitude of ways. The term e-assessment is a broadly-based one, covering a range of activities in which digital technologies are used in assessment. Such activities include the designing and delivery of assessments, marking – by computers, or humans assisted by scanners and online tools – and all processes of reporting, storing and transferring of data associated with public and internal assessments.

If used with skill and imagination, e-assessment can increase the range of what is tested. It can provide evidence of both cognitive and skills-based achievements in ways that are durable and transferable. It can enhance the validity of assessment systems and encourage deeper learning.

*e-Assessment*⁴, in fact, is much more than just an alternative way of doing what we already do. A growing body of evidence indicates that well-designed and well-deployed diagnostic and formative assessments can foster more effective learning for a wider diversity of learners. Assessment is perhaps the best way of identifying the support needs of learners and can instill a desire to further progress if linked to appropriate resources, good quality, timely feedback, and to challenging but stimulating ways of demonstrating understanding and skills. Effective use of technology can make significant contributions here.

Additionally, e-Assessment can support personalisation. Any time, anywhere assessments benefit learners for whom a traditional assessment regime presents difficulties due to distance, disability, illness, or work commitments. On-demand summative assessments, when available, increase participation in learning by enabling learners to advance at a pace and in a way appropriate to them.

Furthermore, some forms of e-assessment may be used at each of the three stages at

⁴ The glossary at the end of the report defines this term.

which a learner's attainment and progress come under review:

- Diagnostic – assessment of a learner's knowledge and skills at the outset of a course.
- Formative – assessment that provides developmental feedback to a learner on his or her current understanding and skills. Formative assessment can also be described as 'assessment for learning' since an assessment that is entered into voluntarily, and on which no final qualification depends, can prompt learners to adjust their own performance.
- Summative – the final assessment of a learner's achievement, usually leading to a formal qualification or certification of a skill. Summative assessment is also referred to as assessment of learning.

Assessment of any kind can be referred to as low, medium or high stakes (McMurrer, 2007), as mentioned above:

- A low-stakes assessment is usually formative, with results recorded locally.
- A medium-stakes assessment is one in which results may be recorded locally and nationally, but is not life changing.
- A high-stakes assessment, however, is one in which the outcomes are of high importance to both centre and candidates, affecting progression to subsequent roles and activities.

2.1 Some important considerations of e-assessment

Some important considerations when rethinking e-assessment practice (JISC, 2007, p.8) refers to:

- Appropriateness
- Timeliness
- Relevance
- Accessibility
- Validity
- Quality of supporting systems

These criteria do not apply solely to e-assessment, but are significant because the application of technology has either proved beneficial, or has prompted a reassessment of

institutional, organisational or academic processes in the following respects:

One way in which the introduction of e-assessment may enhance the quality of the learner's experience is through the closer alignment of assessment with the pedagogic approach used get expanded, in order to include blended and online models, assessment in a similar form becomes increasingly appropriate and relevant.

However, e-assessment should not be viewed in isolation from the practice surrounding its use – for example, the timeliness of assessments and the quality of the feedback provided are key to learners' progress. Technology can assist by supporting on-demand delivery of tests to large numbers of learners, but the real advantage may lie in the immediacy of feedback and opportunities for further learning – for example, regardless the time or the place availability of resources.

Increasingly, a range of methods, both computer-based and computer-assisted, is used in higher education. Online tasks involving peer- and self-assessment and increasing use of e-portfolios and assessment tools within virtual learning environments (VLEs) indicate the diversity of approaches that has been established in this sector. Equally important is the relevance and accessibility of the assessment for the learner: evidence suggests e-assessments can provide assessment experiences that are more authentic – through the use of e-portfolios, reflective diaries, blogs or virtual world scenarios, for example. When objective tests are the most appropriate method, interactive elements and multimedia, or confidence-based marking, can make assessments more valid, accessible and engaging.

PART II

THEORETICAL FRAMEWORK

3 Assessment: Conceptual Approach

Assessment is crucial to learning and teaching. What is assessed defines what is taught and how it is learnt. The process of assessment, in turn, shapes institutional practice and affects a learner's view of the value of engaging in learning. Getting assessment 'right' is essential to the well-being of learners and institutions, and instrumental to the achievement of national strategies for widening participation and eLearning. In general, learning assessments should be observable and measureable. They should be fair, too.

A conceptual model of assessment unit is based on the methods that include the relevant activities (Dureva & Tuparov, 2006), such as:

- Structure, order and type of assessment activities.
- The intend of the activities and reflection of their results in the learner's portfolio:
 - self-assessment;
 - pre- assessment;
 - formative assessment;
 - evaluation and grading.
- The weight of the each assessment activity in the final grade.
- The grade scale.
- The time and duration of the each assessment activity, included in the assessment unit.

However there is still a confusion between evaluation and assessment. To some extent this can be explained by different understandings of evaluation in different countries and by the linguistic confusions between the two processes. However, it goes further than this. There appears to be an over reliance on assessment or achievement as the basis for evaluation. That is not to say that learner achievement is not an evaluation factor. But the assumption that if something has not been assessed it has not been learnt is surely wrong. Furthermore, such an approach ignores informal learning and learning not included in the assessments.

4 Diagnostic, Formative and Summative assessment

Assessment can essentially be divided into three types. Firstly, diagnostic assessment defines the current state of a learner before the learning process starts. It serves as a kind of aptitude test or helps to find the right class and to plan the following. Usually, a collection of students' background information should be done by the registrar's office in order to help elearning manager to understand them. The techniques are:

- Background Knowledge Probe
- Focused Listing
- Misconception/Preconception Check 4) Empty Outlines
- Memory Matrix
- Minute Paper
- Muddiest Point
- Categorizing Grid
- Defining Features Matrix
- Pro and Con Grid

Secondly, formative assessment, which may be administered during the presentation of a course, as a mean of checking on student learning. And thirdly, summative assessment at the end of a period of study, whereby the results are used in order to determine examination outcome. Specially, summative assessment in eLearning environment demands high technical and legal security of e-exams (Hamburger eLmagazin, 2011), since such an assessment determines the final result. Techniques for these types of assessment will be discussed later. However, at this point, a clarification needs to be made about e-exam. The term e-examination does not really content the whole range of use, e.g. the formative use to regulate the learning (assessing the learning). Whereas the term e-assessment has a broader meaning. It contents the whole process of evaluation of a learner, it subsumes examinations alongside with excercises for consolidation of the learned content, as well as the formative evaluation and self-assessment of the students (Schiefner, 2007).

Speaking of terms, it is notable that 'formative' and 'summative' were introduced by Scriven (1967) to describe the various roles of evaluation in curriculum development and instruction. Ebel and Frisbie (1991) defined 'summative assessment' as assessment conducted

at the end of instruction to determine if learning is complete enough to guarantee the moving the learners to the next segment of instruction. Thus, summative assessment is always used to investigate student learning achievement, and is always administered at the end of the teaching process. It is not used to help learners in their study. Rather, it's for the instructor to assess the learner. That type tends to be high-value assessment.

Additionally, Ebel and Frisbie (1991) stated that 'formative assessment' is conducted to monitor the instructional process and to determine whether learning is taking place as planned. However, the purpose of the formative assessment during the teaching process is to illuminate learner difficulties and enhance teacher effectiveness, and is always administered during the teaching process (Angelo & Cross, 1993). This tends to be low-value assessment.

5 Low-Value Assessments

A *low-value assessment*⁵ is one which is low-risk for the learner and which does not involve a lot of points (or no points, or only extra-credit points). A low-value assessment can also be one which may be done again-and-again for aggregated points. The contents of low-value assessments tend to be fairly simple (vs. difficult).

These may be acclimation types of assignments to help learners get used to a learning management system or virtual immersive world—where learners have a kind of treasure hunt in the system. Or these may be syllabus quizzes to ensure that learners have read that electronic document thoroughly.

These types of low-value assessments may be practice assignments. They may be draft assignments that may be revisable for much higher credit. Often, low-value assignments create opportunities for low-risk efforts of learners and opportunities for constructive feedback by the instructor.

⁵ The glossary at the end of the report defines this term.

6 High-Value Assessments

*High-value assessments*⁶ are often those that coalesce plenty of complex learning. These may be research papers, presentations, interviews of experts, field trips and field research, lab experiments, and even group projects. Cumulative midterms and finals tend to be high-value assessments. Comprehensive exams (or oral exams) may also be very high-value ones.

These types of assessments tend to measure student capabilities beyond even the subject matter about which they're being tested. These worth plenty of points, and they may determine entry to or removal from a program of learning.

7 Authentic Assessment

Good teaching practice, especially in higher education and in adult learning, is typified by student and teacher knowledge sharing (McLoughlin, 2002), where students engage in learning that has real world relevance (Lombardi, 2007) that is supported by both the teaching and *authentic assessment*⁷ (Herrington & Oliver, 2010).

The specified elements of design for authentic eLearning according to Lombardi (2007) include having real world relevance where the learning task is based on real life problem solving with a meaningful context for planned learning experiences. There should also be, as Lombardi (2007) suggested, a learning task that provides long-term student engagement with learning that involves a variety of resources and perspectives over a sustained time where collaborations takes place to promote engaging open conversation. Authentic tasks ensure that collaboration is an imperative component of the learning task and learning outcomes. According to Lombardi (2007) reflection is also important as it allows students' time to reflect on their actions while discovering a nexus between classroom theories. Creating assessment questions based on real world is also essential in an eLearning environment.

Assessment questions whether formative or summative are unavoidable in an eLearning environment. Just as people halt during a journey to take a break, similarly eLearning courses have assessment questions to determine how much the learner has been

^{6,7} The glossary at the end of the report defines this term.

able to absorb the content of the course.

8 Why there is a Need for Assessment Questions?

There is a need for assessment questions for various reasons, such as:

- Assessment helps to gauge the effectiveness of the course.
- It helps to assure that the student has mastered the material within the course.
- It encourages and builds confidence in the learners.
- It corrects any misconceptions.
- Acts as a breather.
- It helps to keep track of student's progress.
- It motivates the student to focus on important elements in a course.
- Students/learners also get to know their strength and weakness in terms of performance practice.
- It reinforces learning through feedbacks.

9 Question Types in Assessments

There are several types of questions available in Assessments and within any assessment system question types may vary. For instance, within traditional educational styles, a teacher aims examining students using limited-choice or open-ended questions. Limited-choice questions such as multiple choices are applied to reach lower-level objectives like recalling facts. Open-ended questions like sentence completion, short answers, essays etc. require students to formulate their own answers, which do not have to be pre-determined. It is easy to see that open-ended questions can be used to evaluate higher-level objectives like applying or evaluating assimilated knowledge. Hence, for certain domains like mathematics, physics or programming exercises, limited-choice questions might work for assessing higher-level objective because students have to apply their gained knowledge to complete a question with a pre-determined answer like some mathematical calculation.

Questions, tasks, activities and other methods of eliciting student responses are often

called items in the assessment process. In the computer-based platform, we argue that almost any type of interaction with a user can be considered an assessment item.

Given this definition, there are many ways in which assessment items can be innovative when delivered by computer. One organizational scheme describes innovative features for computer-administered items, such as the technological enhancements of sound, graphics, animation, video or other new media incorporated into the item stem, response options or both (Parshall, Davey, & Pashley, 2000). But there are myriads of other classification possibilities, including how items function. For some innovative formats, students can, for instance, click on graphics, drag or move objects, re-order a series of statements or pictures, or construct a graph or another kind of presentation. The innovation may not be in any single item, but in how the items flow, as in branching through a changing series of items contingent on an examinee's responses.

Much of the literature on item types deals with innovations of the observation – the stimulus and response – that focus on the degree of construction versus selection, or constraint versus openness, in the response format. A number of characteristics are common to most constructed-response and performance formats:

First and perhaps most obvious, these alternative formats require an examinee to supply, develop, perform, or create something. And, typically, these tasks attempt to be more engaging to the examinee than conventional multiple-choice items. They often employ real-world problems that people of a comparable age and peer status may encounter in daily life, such as asking school-age children to calculate from a grocery store purchase, or for high schoolers, to complete a driver's license application or examine an insurance policy. They are generally scored by comparing and contrasting an examinee's responses to some developed criteria, sometimes elucidated in lengthy descriptions called "rubrics". However, the so-called open-ended items cover a multitude of territories, and organizing schemes for the degree of constraint and other measurement aspects regarding items can be helpful.

Others remind assessment developers that items should not require the examinee "*to spend time producing irrelevant data (from the perspective of the measurement goal) or doing irrelevant mental processing*" (Stout, 2002, p. 109). Stout uses the example of a five-minute video clip as a prompt for a five-minute creative essay as perhaps excessive. In answer to this, however, it seems relevant to remember that construct validity is only one goal of assessment.

Other goals may include a better match with classroom activities and objectives, need for more authentic stimuli and responses (and that are engaging as classroom activities), and other aims that may be better satisfied by such media inclusion.

There are also many potential advantages of online assessment to learners. For example, tests are available on demand and at any given time. Furthermore, computerised assessment systems give immediate feedback to the user; therefore users learn by taking the test. However, online assessment systems also have a drawback because students who perceive themselves as possessing poor IT skills may be disadvantaged.

Note that all item types may include new response actions and media inclusion. Thus, by combining intermediate constraint types and varying the response and media inclusion, eLearning instructional designers can create a vast array of innovation assessment approaches and could arguably match assessment needs and evidence for many instructional design objectives.

Media inclusion, simulations, within-item interactivity and data-rich problem-solving in which access to rich resources such as books, resources and references are made available online, are all innovations that can be incorporated in many of the item types discussed below. Examples of media inclusion are numerous and include the multimedia rich National Center for Research on Evaluation, Standards, and Student Testing (CRESST) examples (Chung & Baker, 1997), simulations (Parshall, Spray, Kalohn, & Davey, 2002), and data-rich assessment and tracking of problem-solving paths such as those exemplified in Interactive Multimedia Exercises (IMMEX) (Pellegrino, Chudowsky, & Glaser, 2001).

Below is a description of each type and when each one can be used.

9.1 Multiple Choice

There are two types of *Multiple Choice Questions (MCQ)*:⁸

- Single Correct - there is only one acceptable answer
- Multiple Correct, Multiple Selection - there are several options that are considered "correct" and students may choose more than one. MCQs allow for the answers

⁸ The glossary at the end of the report defines this term.

to be randomized each time the question is displayed to a student so that the correct answer is a different option each time. There can also be a requirement of providing a rationale for student's answer so that there will be verification that the student knows the answer and isn't just guessing.

According to some researchers, ubiquitous multiple-choice testing sometimes encourages *“poor attitudes toward learning and incorrect inferences about its purposes...for example that there is only one right answer, that the right answer resides in the head of the teacher or test maker, and that the job of the student is to get the answer by guessing”* (Bennett, 1993, p. 24).

Movements toward authentic assessment, alternative assessment, performance assessment, dynamic assessment, portfolio systems, constructed response, higher-order assessment and other approaches favoring richer assessment tasks are often based on consequential validity arguments about deleterious effects on teaching and learning of narrow assessments in the classroom (Osterlind, 1998).

Some cognitive theorists argue that the multiple-choice format presumes, often without sufficient basis, that complex skills can be decomposed and decontextualized. Moreover, some critics maintain that in practice, this format over-relies on well-structured problems with algorithmic solutions and that in theory, it builds on a view of learning that knowledge is additive rather than integrative of developing knowledge structures (Glaser, 1988, 1991; Resnick & Resnick, 1992; Shepard, 1991a, 1991b). But multiple choice type of question have clear advantages when used with large number of users (1.000 or more) (Khom, 2006).

Besides the weakness that was outlined, MCQ are readily scorable and less susceptible to guessing in comparison to true/false and alternate choice formats since more answer choices are available. However, overuse or exclusive use of multiple choice has been criticized as decontextualizing and decomposing learning, encouraging poor habits of mind and reducing the richness of the instructional experience by teaching to selection-only test formats (Resnick & Resnick, 1992). Additionally, limiting computer-based approaches to standard multiple choice items may not optimize the assessment potential of the technology platform, where more construction can be readily achieved.

Innovations in the multiple-choice category for online settings can include new response actions not common in paper-and-pencil settings, such as clicking on an area of a

graphical image, and can also include new media, such as distractors that are sound clips (Parshall, Spray, Kalohn, & Davey, 2002).

9.2 True/False

*True/False (TF) questions*⁹ ask the student to determine if a statement is true or false. Research on the TF format generally supports its use, except for the detriment of unfamiliarity by item writers (Haladyna, 1994a). According to Haladyna (1994b), TF and conventional MCQ were compared in a medical testing situation that found TF yielded more reliable scores. However, conventional MCQ was more highly correlated with complex measures of competence and, at least in the examples studied, TF items seemed to be a measure of more basic knowledge.

The yes/no explanation format, goes one step beyond being a set or bundle of true/false questions. It requires a two-step reasoning process involving identifying which alternative is correct and then recognizing why it is correct, for each true/false pair (McDonald, 2002).

McDonald (2002) cites this item type, if well-written, as tending to have higher discrimination indices, or more readily distinguishing between students of higher and lower abilities, than conventional multiple choice. It is challenging to write well, though, as each set of explanations must be plausible for each pair choice. Otherwise students may have clues as to which answer is correct just based on logical reasoning and elimination of unlikely explanations.

9.3 Essay

Essay questions provide students with two ways for submitting their answers.

- A text box with a Content Editor

⁹ The glossary at the end of the report defines this term.

- The ability to attach a file to the question

The instructor decides to allow one or both of these options. *This question type must be manually graded.*

9.4 Matching/ Reordering/Rearrangement

*Matching type question*¹⁰ require matching one term, phrase or image to another by the student.

Haladyna (1994b) says there is little research to report on unique features for the matching item format. He recommends that such items should continue to be used as a variation of conventional multiple-choice since they are easy to construct and administer; they are efficient in space, as options do not have to be repeated, and in test-taker time. Besides the lack of research and theory, other limitations for the matching type come with item-writing traps that are easy to fall into, including nonhomogeneous options, such as mixing sets of things, people and places, or providing equal numbers of items and options, both of which make guessing easier and can bring test-taking skills into play as a nuisance, or unwanted, dimension of performance

9.5 Completion

The *completion question type*¹¹ involves items in which not all the components for the answer are supplied in advance, or in which the entire problem-space is contained in the set of answers. The assessment literature tends to treat all these formats as innovative types, or alternative assessments, to a greater or lesser degree.

The completion category asks respondents to finish an incomplete stimulus. Item types include single numerical constructed items, short-answer and sentence completion, Cloze-procedure, and matrix completion problems (Embretson, 2002). Much recent work on the completion format focuses on techniques of automatic scoring, which is not treated in this paper. A good treatment of this topic is available in a book on automated scoring, which includes a review chapter across numerous approaches (Scalise & Gifford, 2006).

^{10,11} The glossary at the end of the report defines this term.

This item format was once assumed in the literature to be best for low task complexity (Parshall, 2002), but it seems perhaps an unnecessary limitation as items demanding complex problem-solving, strategy selection and solution construction can result in single, well-defined numerical answers. This is how the item type is often used in the classroom, although with the specification that students show their work so that the problem-solving process is more clearly elucidated for partial credit scoring and learning intervention.

An disadvantage in using constructed response is that sometimes these items are usually omitted by respondents. In the National Assessment of Educational Progress (NAEP), for instance, constructed-response tasks are omitted far more often than multiple-choice questions (Bennett, 1993).

9.6 Presentation/Portfolio

Falling into the least constrained or “*presentation/portfolio*¹²” category are a wide variety of complex performances that include such activities as projects, portfolios, demonstrations, experiments, fine art performances, and medical diagnoses or other professional practicum as well as teaching and extended group activities, discussions, and interviews. There are large bodies of work on such assessments, which some describe as performance assessments, although this term has multiple meanings and can refer to item types in more constrained categories, as well (Gronlund, 2003).

For assessment systems with considerable sophistication in the available scoring algorithms, is sometimes possible to generate computer-based scoring for some aspects of assessment tasks in the presentation/portfolio category; however, there are many challenges to the validity of these scores and often human scoring or comparable human scoring is desirable.

Secondly, computers can create opportunities for innovation that involve group activity at distributed locations or at asynchronous times. Peer assessment is an example and so are activities that call for engaging in a community of practice. Knowledge is often embedded in particular social and cultural contexts. Although there is much need for assessment that involves students engaged with other students, there is limited knowledge of best practices

¹² The glossary at the end of the report defines this term.

for group assessment (Pellegrino, Chudowsky, & Glaser, 2001). Technology may offer some new tools, like those (Koch, 2013):

- **Discussion fora:** are often used in the context of eLearning for discourses and cooperative activities. A discussion forum as a form of e-assessment can either help to add a certain authority to discussions and to admonish students to online communication, or it is used for pedagogic- didactical reasons, if "reflection" and "discussing competency" is the educational objective of a learning unit. For example, in an evaluation students/learners have to launch or advance a discussion about a certain topic in a forum for example.
- **Weblogs:** study diaries are used in educational science for quite some time. They can also be implemented on the computer in form of weblogs. Weblogs are frequently updated web pages, whose articles are arranged chronologically in categories, at which the latest article always comes first. Weblogs can be tools for peer collaboration, public or private conversation, reflective online portfolios and life long learning spaces. They can be used as instruments of active learning (active knowledge acquisition/sense-making, gradual improvement through continuous preoccupation with the learning content, self-directed learning). But weblogs are also suitable for group work by sharing one`s own thoughts and discourse through commenting. Furthermore, weblogs reinforce the responsibility of an individual as the contributions to the weblog are personalized. A weblog enables the teacher to give an individual, blog-related feedback to a learner.
- **Wikis¹³:** wikis are asynchronous, editable web pages that enable collective text-writing, the compilation of knowledge and project management. A great benefit of wikis in the context of e- assessment consists in picturing the learning process of the learners. Due to the version-control, i.e. the possibility of tracing back single steps of the revision of an individual or in team work created article, the singular thinking and writing steps of the learners are visible and documented.
- **E-portfolios¹⁴:** An e-portfolio is a deliberate choice of electronic documents and certificates. It can be used in different settings and can be used in different functions.

¹³ The glossary at the end of the report defines this term.

^{14, 15} The glossary at the end of the report defines this term.

Above all it serves the documentation of knowledge, skills and the learning process. Furthermore it is a record of the learner`s personal development, application processes, performance evaluation. Expertise can be proved by e-portfolios if they are used as complements and consolidations of learning units. E-portfolios can advance interdisciplinary and metacognitive competencies, as a deliberate choice of the elements of a portfolio conditions a reflection and an abstraction of the single items of work achievements and artifacts. Thereby the main function of e-portfolios lies in self-reflection and self-improvement of the learner (Schiefner, 2007).

- **Peer Assessments**¹⁵: Shared learning spaces where peers may post and share their work enable peer assessments. The rubrics used for the assessments may be those that are instructor-created, student-created, or instructor-and-student created. The peer assessments are generally anonymous, and the focus is to develop both the learners receiving the constructive critique and the learners creating the constructive critiques.
- **The observation of students' behavior in chat room**: eLearning requires students to interact with other students in the chat room. If chance allows, a tutor or instructor should observe students' behavior in order to find out:
 - 1) Who chats with whom?
 - 2) What do they chat about?
 - 3) Is the content related to the course?
 - 4) Does the content lead to understand critical thinking?
 - 5) If there are more than 3 persons, who leads the group?
- **Further forms of e-assessment**: webquests, video-exams, termpapers, badges, case-studies, etc.

9.7 Question Types in Assessments: Conclusions

Assessment task design is a rich and complex arena in the rapidly emerging field of

computer-based assessment and involves many considerations, including interactivity, the flow between items, assessment assembly specifications, and considerations of item feedback and learning interventions intermingled in the item flow.

Many other innovative item formats can be derived from combinations or adjustments of the example types, and item formats across types can be widely varied depending on the domain to be measured and the inclusion of new media such as interactive graphics, audio, video, animation and simulation.

Mislevy (1996) makes the point that if the information provided by an innovative item type is no better than the one provided by conventional multiplechoice, then the innovation seems pointless. In other words, innovations must be justified by providing something beyond what is available through standard formats. For innovative assessment questions and tasks, what this “something” is might take many forms, from increasing predictive validity, to improving classroom effects, or to providing better metacognitive interventions by increasing the ability to diagnose paths to competency rather than simply ranking students.

10 Different (or Impossible) Online Assessments and Possible Mitigation

Traditionally, it is thought that some types of assessments should be done in a face-to-face setting. Orals are usually held in a face-to-face way with the doctoral committee. This is mitigated with real-time video-to-video assessment to mitigate for the distance. (Web conferencing is not at the quality where this has been used commonly as a substitute.)

Some difficult laboratory experiences require direct human presence because there are no simulations for many of the complex machines that are used. Some mitigation have been to use local experts to oversee some of this work, but this involves a lot of paperwork to ensure the quality of the learning and assessment.

11 Methods for Formulating the Right Mix of Assessments

The subject matter experts have wide discretion in formulating the right mix of assessments for learners. Some basic principles follow:

- Offer a variety of assessments to accommodate those with various learning styles

- Begin with low-value assessments to help learners acclimate to the technologies and the learning domain
- Spell out all expectations for all assessments
- Offer work examples where helpful for divergent types of assessments (Make sure to have student permissions before posting their work.)
- Make sure that there is sufficient time for learners to prepare for high-value assessments
- Make sure to have accessibility mitigation for those who have special needs (Brown, 2002, pp. 61 – 68).
- Have a walk-through of a course to make sure it all makes sense. Keep channels of communications open with learners, so that mistakes may be corrected, and changes may be made to enhance the value of the assessments for both formative and summative purposes.

12 Feedback in formative assessment

An important function of assessment is providing students with ‘continuous feedback’, meaning that opportunities for feedback should occur continuously, but not intrusively, as a part of instruction (Williams, 2004; Bransford *et al.*, 2000). Bell and Cowie (2001) concluded that assessment –mainly the formative one- is increasingly being used to refer only to assessment that provides feedback to students (and teachers) about learning occurring during the period of instruction and learning, and not after that. Feedback from formative assessment is beneficial in the adjustment of teaching strategies and application of appropriate remedial techniques. Bell and Cowie (*ibid.*) suggested that the teacher gather assessment information (feedback) about student learning from formative assessment, and then respond to promote further learning, and then shape and improve student competence.

Feedback in formative assessment can uncover weaknesses requiring reinforcement and is seen as an essential component of the formative assessment interaction, where the intention is to support learning (Sadler, 1989; Perrenoud, 1998). Teachers should make use of formative assessment to give students feedback at an appropriate point in the learning

process (Brown & Knight, 1994) because formative assessment will be of little help to learners if teachers do not allow time for feedback. Messages from the feedback of formative assessment should tell learners what must be done, thus improving learning effectiveness. In traditional classroom teaching, teacher feedback to learners is limited (Bransford *et al.*, 2000, pp. 140–141). Typical teachers give just one summative assessment, like transcripts or assignment scores. After grades are given, students move on to the next topic and work for another grade. In other words, learning is driven by reading and making grades. Such learning is not effective. Instead, the truly meaningful feedback that teachers give learners should derive from formative assessment. By using meaningful feedback, students can improve weaknesses in learning and thinking, increase and transfer learning, and value opportunities to revise (Barron *et al.* 1998; Black & Wiliam 1998; Bransford *et al.* 2000, p. 141).

Thus, it is important for teachers to provide learners with opportunities for receiving ‘timely feedback’ and ‘repeating the test’. In order to help learners and tutors to achieve such a good feedback Haynes et el. (2004) has suggested a strategy to encourage students to use feedback. First of all, Haynes et al. (2004) suggests good feedback should:

- Be based on the criteria for the assessment in question
- Give credit for what the student has learnt
- Be designed in order to give information on the quality of their work and its strengths and weakness
- Concentrate on the student’s achievements, emphasizing their strengths and how they have improved
- Be positive and directed towards helping the learner to improve their performance;
- Be motivational
- Be not overwhelming
- Encourage the students to reflect on their own work and work towards managing their own learning.

When students send in a piece of work, Haynes et al. (2004) suggests e-teachers to encourage their students to add a separate sheet indicating:

- What they are pleased with
- What they think they have done well (or tried to do well)

- Which things were difficult and they are not too happy about, and about which they would particularly like comments
- Their own assessment of the piece of work
- Any other comments

After receiving the work back, Haynes et al. (2004) suggests that students should do the followings:

- Ask for further explanation muddled answers
- Having read the grade to put the assignment on one side for a few days
- When they are ready, to skim through the actual comments
- Then they should read through their assignment, stopping to read the comments in the appropriate places. They should mark in a different colored pen their responses to what is written – anything they agree or disagree with, or anything they don't understand
- They should now think about the grade again; if they are still unhappy, they should email their tutor
- On a positive basis, they should look at the comments once more, and pick out two or three main points – pieces of advice, errors, things to remember: points which they could bear in mind when they write their next piece of assessed work.

13 SOME MORE BASIC DIDACTICAL ASPECTS OF ASSESMENT ACTIVITIES

13.1 Learning objectives

The *learning objective*¹⁶ describes the level of knowledge, skills, competences and other characteristics that any learner should achieve in the process of learning. Learning objectives drive the whole educational process. They explore what content should be included in the learning materials and at what level of the cognition of the learning material should be mastered. The learning objectives have to be measurable.

For the description of learning objectives two basic domains, namely knowledge domain and cognitive domain could be used. The knowledge domain describes the concepts,

^{16,17} The glossary at the end of the report defines this term.

facts, procedure, and appearance from concrete learned material. Usually the cognitive domain is described according to some well known taxonomy. One popular and useful one is the Bloom's cognitive taxonomy, which created in 1956, based on six cognitive levels- Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation.

13.2 Criteria for success

The *criteria for success*¹⁷ or so called "cut off level" could be described on the base of expert evaluation. The experts define the minimal level of mastering for each learning objective in percentages. The criteria for success could be described in a matrix of cut off levels according to the expert evaluation and relevant statistical method for expert evaluation.

PART II

LITERATURE REVIEW

Here we present a segment of a published body of knowledge through summary of prior research studies, reviews of literature and theoretical articles. We group research studies and other types of literature (theroretical articles, case studies, etc.) according to country per partner of Criton programme. Only Sweden didn't condribute to this task.

14 The case of Germany: Lessons Learned - 7 years of e-assessment

Annette Koch (AuB) conducts a literature review for the case of e-assessment in Germany. At the university of Bremen there are summative e-assessments taking place since 2004. According to Bücking (2011) the most important long-term conclusions are:

1.High-quality e-assessments require didactical consultation and training as well as the reinvestment of the time (which was saved during the phase of correction) into the creation of good questions and into the creation of open questions according to the educational objectives.

The validity of exams does not depend on the form of an exam. Whether it takes place electronic or handwritten, whether the questions are presented as closed-ended or open-ended assessments, none of it indicates something anything about the quality of the assessment. The widespread prejudice that says that exams on the computer only include MCQs and test factual knowledge is easy to disprove. On the one hand, e-assessments can include a great variety of examination forms (more than it is possible in handwritten exams), on the other hand, there are also MCQs which are able to test comprehension and transfer, even though a very challenging task for the authors is to create these kind of MCQs. Whereas individual assessment and creative achievements require open testing formats and thereby produce a higher postprocessing effort. Practical experience shows that the above listed

prejudices often become reality because of the relatively little temporal investment for the creation of adequate questions, the abandonment of open questions for capacity constraints and too little didactical know-how in the drawing up of good exams.

2. Transparency in matters of educational learning objectives and examination formats (e.g. by offering a mock exam) is even more important for e-assessments than it is for other examination forms.

In e-exams with exclusively closed-ended questions most students expect to be asked factual knowledge and therefore learn some topics by heart. Interviews with teachers and students showed that even if transfer or comprehension questions are asked, many students cling to their expectations and only prepare by learning special themes by heart and afterwards they complain about unfair or incomprehensible questions.

3. Formative assessments should be promoted and further developed in their efficiency. At the same time, learning institutions must work towards the creation of high-quality exams, designed to build competencies.

In many cases formative assessments make more sense than summative assessments. It is well-known that the efficiency of summative examinations is very low for lasting competency development. Although e-assessments comprise much more than summative examinations, large numbers of students and the higher effort concomitant the formative assessments hardens the replacement of traditional forms of assessment.

4. A test center contributes considerably to the establishment of safe and efficient examination procedures.

5(a). To accommodate a high demand for e-examinations a transition to flexible examination periods is indispensable.

This is often impeded because of rigid curricular structures.

5(b). Sustainability and high-quality assistance/supervision is only attainable through perpetuation and financial protection of the e-assessment service.

The maximum number of e-assessments per semester is not limited by the availability of assessment spaces but by the capacity of the supervision through the eLearning team.

6. Legal uncertainty and technical risks are overrated. This is made clear on the example of the "Bremer Modell" where there have been 47.000 successful performed e-examinations and without any legal proceedings or assessment cancellation until today.

7. Cutting down on expenses is greater in the field of open-ended forms of e-assessment than in the case of closed-ended forms of e-assessment.

That means a relief through e-assessments especially for teachers of philosophy, history, philology, arts, social sciences, etc. A possible reason for the saving of time and cost in open-ended e-assessments is that unreadable handwriting is eliminated through computerized examinations.

These unexpected results came to light in the summer of 2009, during comparative measurements for the phase of evaluation. The aim was to calculate the expectable saving of expenses (time and money) for the tutors. The working hypothesis was that e-assessments develop above all their full potential in the field of standardized and thereby automated evaluable exams with closed-ended questions. The measurements came to the conclusion above (Hamburger eLmagazin, 2011).

15 The case of Austria: e-Assessment practices

The research about E-Assessment practices in Austria resulted in five main texts about e-assessment practices. The keywords used in the research phase were the German equivalents for E-Assessment, best practice in e-assessment, evaluation of learning outcomes, and e-testing. The search was carried out from Katharina Resch (deBerater) at the period from March to May 2013 in google scholar, main libraries and publishing houses' websites.

The search did not produce vast results, since there are a number of publications on eLearning in Austria, but almost none of them focus on e-assessment. The University of Vienna Library for example has 236 books on eLearning, but not anything about assessing eLearning.

The main term „E-Assessment“ is used in the German language used as a term for two different processes: on the one hand for assessing learning outcomes and performance of online courses and studies and for recruiting processes in organizations, who work with e-assessment tools in human resource management on the other hand.

The five main texts are summarized in the below Table 1:

No.	Author(s)	Title	Place	Publisher	Year	Form
1	Andreas Raith	E-Testing im österreichischen Hochschulbereich – eine Chance?	Krems	Zentrum für Bildung und Medien. Abteilung Telekommunikation, Information und Medien, Donau-Universität Krems	2004	Book
2	Susanne Gruttmann & Claus Usener	Prüfen mit Computer und Internet. Didaktik, Methodik und Organisation von E-Assessment	Graz	In: Schön, Sandra; Ebner, Martin (ed.). Lehrbuch für Lernen und Lehren mit Technologien (L3T), Universität Graz	2011	Article
3	Ulf-Daniel Ehlers	Qualitätssicherung im E-Learning. Veränderung durch derzeitige Technologien und Konzepte	Graz	In: Schön, Sandra; Ebner, Martin (ed.). Lehrbuch für Lernen und Lehren mit Technologien (L3T), Universität Graz	2011	Article
4	Walter Khom	E-Testing. Die konsequente Fortführung von eLearning	Graz	bit media, Konferenzbeitrag am Virtual Campus FH Joanneum	2006	Conference Proceedings
5	Patrick Hoitsch	Entwicklung und Einsatz von e-Testing Szenarien	München	GRIN Verlag GmbH	2008	Book

Table 1: E-assessment practices in Austria – Literature review

Based to Katharina Resch (dieBerater) and her literature review, Andreas Raith (2004) in his book titled *“Does E-Testing have a chance at university level?”* (*“E-Testing im österreichischen Hochschulbereich – eine Chance?”*) begins with an overview of traditional ways of testing and examining and then contrasts this with E-testing. Traditional ways of testing were written as well as oral exams, assignments during a time of some weeks and

months resulting in a cumulated grade (process grading) and portfolio testing. E-testing is different to these and is simply defined as computer-mediated testing via internet or intranet. Self-learning is excluded from the book.

E-testing involved a lot of competences of the teacher, to design adequate questions and assignment types. Adequate types are: single choice, multiple choice, true/false questions, short answer questions, keywords, matching, quizzes, and essays. Other competences of the teacher should be to administrate the exams technically, design the exams, and analyse them.

Advantages and disadvantages of e-testing are also discussed. The positive points are that they do not depend on time and space, that they can be analysed quickly, that their assessment is transparent and easily standardizable. The negative ones are accessibility problems of students to the computer, technical problems using the computer, and open legal questions.

Then the costs of using traditional testing methods and e-testing methods are compared. If a written exam with paper/pencil method is compared to an e-test with 1.000 pupils each, the results indicate that a traditional examination need 24 hours time in preparation, implementation, and analysis and that the e-test in comparison only needs 4 hours for the same number of pupils. The estimated costs of a traditional exam are calculated with app. 10.000 €, the e-test only with 2.200 €.

Next resource comes by Susanne Gruttmann & Claus Usener (2011). According to this article by the title *“Assessments with the computer and the internet – Didactics, Methods, and Organizing E-Assessments”* (*“Prüfen mit Computer und Internet - Didaktik, Methodik und Organisation von E-Assessment”*) e-assessments have specific requirements regarding didactics, methods, administration, and technical requirements compared to traditional assessments. E-assessments aim at promoting learning and selecting those who do not fulfil learning outcomes.

Two types of e-assessments are discussed in the text: Summative assessments want to measure if learners have reached certain learning outcomes. Formative assessments aim at measuring learning progress in several steps or times.

The article also differs between two adequate types of e-assessment questions: convergent and divergent assignments. Convergent assignments have a fixed number of answers, like multiple choice or true/false answers, and divergent assignments have an open number of answers which involve background knowledge of the learner, like short texts or essays. E-assessment makes a lot of sense for convergent assignments, but is practically not used very much for divergent assignments.

Gruttmann & Usener (ibid.) give an overview of administrative tasks for e-assessment, including preparation like producing participant lists, prevention of manipulation, or filing old assignments. The text ends with a summary and potential and barriers to e-assessments.

On the other hand, Ulf-Daniel Ehlers (2011) at his article by the title *“Quality assurance in ELearning. Changes based on current technologies and concepts”* (*“Qualitätssicherung im ELearning. Veränderung durch derzeitige Technologien und Konzepte”*) talks about learners 2.0 who learn with e-portfolios during the course of a semester and exchange knowledge about a subject with other learners, but also with teachers and experts. The learner generation “ne(x)t generation” learns online using all available spaces.

Ehlers (ibid.) wants to know how the quality of learning in web 2.0 can be measured and assured. Assessing eLearning processes requires a high level of autonomy of the learner and measuring these processes has to concentrate on learning outcomes and individual performance. The text shows the differences in quality assessments in traditional and new learning processes. In new processes the quality of a learning outcome is assessed by peers or learners themselves in self-evaluation and not by experts any more. Contents of subjects are made by learners themselves and no longer by teachers. While traditional assessment methods are about what learners have learned, the new ones are about participation.

Teachers will play a new role, not the one of assessing learning processes, but of reflecting them together with learners. Learners are in the new role of not receiving learning materials, but of producing them.

Walter Khom (2006) in his conference paper by the title *“E-Testing. The consistent continuation of E-Learning”* (*“E-Testing. Die konsequente Fortführung von e-Learning”*) points

out that measuring skills is becoming more and more important. E-testing is available everywhere and it is an objective form of assessment. It involves a number of roles and staff members: the learner, the teacher, the educational controller, the educational administration, and the educational planner of the organisation. As Gruttmann & Usener (2011) mentioned, e-assessment involves different types of assignments, especially multiple choice and test as well as simulation. Both forms have clear advantages when used with large number of users (1.000 or more).

Khom (ibid.) also gives three examples of organisations and how they use e-assessment and he shows the process logics of e-assessments: the administration sends out a date for an e-test, then learners can register for it, then modules or specific e-tests are allocated to the learners, then the e-test starts and is completed, and is followed by an analysis. The teacher has to end the e-test after it is finished. The text ends with the advantages of e-testing.

Similarly, Patrick Hoitsch (2008) by the book *“Development and Usage of E-Testing Scenarios”* (*“Entwicklung und Einsatz von e-Testing Szenarien”*) discusses e-testing. He recognizes that electronic assignments are integral part of modern education and training and have a positive influence on learning and skills acquisition. E-Testing can also help to improve organizational and administrative processes of testing in colleges or universities. Based on the current discussion about teaching quality in Austria, this book aims at introducing technical and didactical requirements for using E-testing in Austria and to give recommendations for the future. An empirical study at the College *“Fachhochschule CAMPUS 02 Graz”* is included.

The first part of the book gives an overview of international and national educational activities describing theories and strategies of learning, focusing on new media integration and electronic assignments and testing. After looking at advantages and disadvantages of e-testing, different aspects like types of questions in e-assignments, or types of tests etc. are discussed.

Then e-testing tools of moodle and dynamic power trainer are introduced and their functionalities are compared. Five different scenarios of integrating e-testing into learning

processes are discussed in the practical part of the text. The text ends with recommendations for teachers who use e-testing.

16 The case of Greece: Alternative assessment methods, evaluation and quality assurance in Greece – Literature review

Literature review in Greece has been conducted from Evagelia Bufardea (CTI) and Adamantia Spanaka (HOU). There is a variety of resources, mainly from Conference Proceedings in Open and Distance Learning. The following Table 2 shows the most relevant papers about e-assessment and evaluation in distance learning.

No.	Author(s)	Title	Place	Publisher	Year	Form
1	Antonis Lionarakis & Adamantia Spanaka	Formative evaluation in distance learning	Athens	Open Education - The Journal for Open and Distance Education and Educational Technology	2010	Article
2	Evagelia Gouli, Agoritsa Gogoulou & Maria Grigoriadou	Supporting alternative assessment methods through PECASSE environment	Syros	5 th Conference of ICT in education	2009	Conference Proceedings

Table 2 (Continued on next page): Alternative assessment methods, evaluation and quality assurance in Greece – Literature review

No.	Author(s)	Title	Place	Publisher	Year	Form
3	Christos E.	Integrating E-learning	USA	World Academy of	2006	Article

	Alexakos, Konstantinos C. Giotopoulos, Eleni J. Thermogianni, Grigorios N. Beligiannis and Spiridon D. Likiothanassis	Environments with Computational Intelligence Assessment Agents		Science, Engineering and Technology		
4	A. Giannakulopulos & M. Meimaris	Evaluation of courses via the web and proposals for applications	Athens	4 th International Conference in Open and Distance Learning (ICODL)	2007	Conference Proceeding
5	Sylvie Ioakimidou, Antonis Lionarakis	The quality assurance in teaching and learning processes in distance education: a review	Loutraki	6 th ICDL	2011	Conference Proceeding
6	Sofia Papadimitriou, Spyros Papadakis, Antonis Lionarakis, Achilleas Kameas	A proposal for the use of Learning Activity Management System (LAMS) to support the work of Tutors in HOU	Loutraki	6 th ICDL	2011	Conference Proceeding
7	Georgios-Aristeides Papathanasiou, Evangelia Manousou	E-portfolio as tool implementation supplementary school distance learning	6 th ICDL	2011	Conference Proceeding	6 th ICDL

Table 2: Alternative assessment methods, evaluation and quality assurance in Greece – Literature review

There are theoretical papers and original research articles. More precisely, the first paper is a theoretical one. Authors Antonis Lionarakis and Adamantia Spanaka (2010) in “*Formative evaluation in distance learning*” aim to select the necessity for formative evaluation in higher education institutions, such as in the Hellenic Open University. They

describe the content of quality assurance and evaluation process through the international bibliography, leading to proposals for the evaluation in Hellenic Open University.

Evagelia Gouli, Agoritsa Gogoulou & Maria Grigoriadou (2009) in the paper *“Supporting alternative assessment methods through PECASSE environment”*, they present the basic functions of the e-learning environment PECASSE, which was designed and implemented in order to support alternative assessment methods, such as self-assessment, the peer review and collaborative assessment. Distinguishing features PECASSE concerning the support of all three methods of evaluation and possible combinations, the possibility of collaboration of students with multiple ways to form groups of authors and / or assessors following alternative strategies and to define the shape evaluation following alternative approaches target the active involvement of students and guide them in the evaluation phase.

Christos E. Alexakos, Konstantinos C. Giotopoulos, Eleni J. Thermogianni, Grigorios N. Beligiannis and Spiridon D. Likothanassis (2006) in their contribution *“Integrating E-learning Environments with Computational Intelligence Assessment Agents”* present an innovative platform that integrates intelligent agents in legacy eLearning environments. This article introduces the design and development of a scalable and interoperable integration platform supporting various assessment agents for eLearning environments. The agents are implemented in order to provide intelligent assessment services to computational intelligent techniques such as Bayesian Networks and Genetic Algorithms. The utilization of new and emerging technologies like web services allows integrating the provided services to any web based legacy eLearning environment.

A. Giannakulopulos & M. Meimaris (2007) examine *“Evaluation of courses via the web and proposals for applications”*. The paper examines the anticipated differences within the evaluation procedure of courses and tutors as a result of the use of ICT and proposes specific examples of applications, in which the evaluation is substantiated via the web thus achieving the desired ease of use and quickness. Apart from the technical matters, the discussion focuses on the anonymity of the evaluators and the temporal parameter. Though the technical restrictions are not overlooked within the conclusions, the proposed modifications

clearly support the enrichment of the platforms with modules for the evaluation of courses and tutors in order to result in a more effective process.

Another theoretical paper comes from Sylvie Ioakimidou and Antonis Lionarakis (2011) by the title *“The quality assurance in teaching and learning processes in distance education: a review”*. This paper attempts a review of the literature on quality assurance in open distance teaching and learning processes mainly in universities during the last two decades. There is an effort to bring forth some elements about the aims; the kind of research and the methodology used; some matters or problems faced so far; some tendencies and orientations to future research. Influenced by the management field the quality assurance in education has gradually become an issue of great importance and interest. Teaching and learning in the core of the educational process cannot but be affected. This research review could be taken under consideration so far or it could work as a kind of motivation for future research on this field focusing on issues like a wide agreement on objective and reliable measurement of quality in teaching and learning processes in an open distance learning environment.

Sofia Papadimitriou, Spyros Papadakis, Antonis Lionarakis, Achilleas Kameas (2011) make *“A proposal for the use of Learning Activity Management System (LAMS) to support the work of Tutors in HOU”*. Based to that, the integration of the Information and Communication Technologies (ICT) has changed significantly the field of education. Due to the adoption of new technologies, eLearning has been emerged and developed. As a result, distance learning has transformed and new possibilities have appeared. It is remarkable that distance learning became and considered as a scout of the new era in education and contributed to the quality of education. Supporting tutors by means of advanced learning technologies in distance education is a significant contribution for their effective role in organizational, social and educational context.

The Learning Activity Management System (LAMS) is a new proposal to the learning design standard. In this paper, we propose the LAMS exploitation at the Hellenic Open University (HOU) and illustrate a guide to design and develop a sequence of learning activities on *“Preparing and producing assignment”*. The preparation of the assignments by students is

the primary mode of the comprehension of educational materials and an important dimension of teaching students in distance education context, where they are the main responsible for their learning. This proposal highlights the advantages, requirements and constraints concerning the development of learning sequences specifically designed by tutors themselves. They follow the principles of adult learning and development of distance education material, oriented to different educational goals and individual needs of their learners.

Finally, the paper *“E-portfolio as tool implementation supplementary school distance learning”* by Georgios-Aristeides Papathanasiou, Evangelia Manousou (2011) argue that the increasing need of an improved education and the upliftment of standards in learning, is leading to a development and use of digital technology which enables the creation and the implementation of tools supporting and strengthening the learning process in school distance education. They present the digital student file (PS.F.M.), as an innovative tool applied to the secondary education in our country in the process of implementation of the institution's School Career Guidance (SEP). This study aims at investigating the creation method of a digital student file, and its development within the framework of school additional distance education in the discipline of School Career Guidance in Secondary Schools. This investigation was carried out through an extensive literature review on Greek and foreign literature within the theoretical framework and its applications, the student work folder, the digital student file, the development of an application methodology, and the methodology of educational research.

17 The case of Lithuania: Literature review of Assessment in conventional education and in eLearning environments

Literature review in Lithuania has been conducted from Gileta Kierené et al. (Soros

International House) timely on May 2013. The following list shows the most relevant articles and reports about assessment in conventional education and in eLearning environments:

1. Gudauskas R., Simasius R. (2008). *The Development of eServices in an Enlarged EU: eLearning in Lithuania*. Luxembourg: Office for Official Publications of the European Communities EUR – Scientific and Technical Research series – ISSN 1018-5593. Retrieved from The Institute for Prospective Technological Studies (IPTS) website: <http://ftp.jrc.es/EURdoc/JRC47549.pdf>
2. Ciuzas R. (2011). Student assessment: Are we ready to shift from assessment of learning to assessment for learning? *Social Sciences*, 71(1), 73-79. doi: 10.5755/j01.ss.71.1.385
3. Vitienė N., Mičiulienė R. (2008). Application of criteria-referenced assessment and qualitative feedback to develop foreign language speaking skills in the context of e-teaching/learning. *Quality of Higher Education*, 5, 152-168, 170-171. Retrieved from http://skc.vdu.lt/downloads/zurnalo_arch/amk_5/qhe_2008_152_171.pdf
4. Kaklauskas A., Zavadskas E.K., Pruskus V., Vlasenko A., Seniut M., Kaklauskas G., Matuliauskaite A., Gribniak V. (2010). Biometric and intelligent self-assessment of student progress system. *Computers & Education*, 55, 821-833. Retrieved from <http://www.sciencedirect.com/science/journal/03601315/55/2>
5. Rimkuvienė S., Lepkova N., Krutinis M. (2010). Results of three research works on e-learning with a special emphasis on the change of economic conditions. *Modern Building Materials, Structures and Techniques*. Selected papers of the 10th International Conference (pp.506–511). Vilnius, Lithuania. Retrieved from http://dspace.vgtu.lt/bitstream/1/517/1/0506-0511_rimkuvienė_lepkova_et_al.pdf

The first paper from the above list, has been written by Gudauskas and Simasius (2008), based on a research by Institute for Prospective Technological Studies (IPTS). IPTS is one of the seven research institutes that make up the Joint Research Centre of the European Commission launched a project to support eGovernment, eHealth and eLearning policy. The research, which was carried out in 2005, focused on the three application areas in the ten New Member States (among them Lithuania) that joined the European Union in 2004, in order to build up a picture of their current status and developments in the field, the most important opportunities and challenges they face, the lessons other member states may learn from them, and the related policy options.

The specific report by Gudauskas and Simasius (2008) was produced by the Lithuanian Free Market Institute, the consortium member from Lithuania, and it presents the results of the research on eLearning in Lithuania. First, the report describes Lithuania's educational system and the role played by eLearning in it. Then, the major technical, economic, political, ethical and socio-cultural factors of eLearning developments. These provide the basis for the identification and discussion of policy options to address the major challenges and to suggest research and development issues for facing the needs of the country.

Gudauskas and Simasius underline that the main thing lacking in Lithuania is the motivation to deliver eLearning services that create **added value** for users and learners. Main policy measures to foster eLearning to the optimal level should target the system of motivation of public education institutions, and of their personnel. The structure and motivation of public education institutions is a much broader aspect than just eLearning. However, this aspect is of primary importance to the prospect of eLearning and of integration of ICT into the conventional system of education. The traditional educational institutions should be encouraged to use more ICT in the study process and to apply the range of eLearning methods.

Based on this report, the most relevant eLearning-related issues of public debate and the challenges for the future in Lithuania are:

- Centralization vs. decentralization in eLearning developments;
- The question of open source vs. commercial software usage in public eLearning;

- Standardisation, integration and interoperability of different learning resources and services, at both national and EU levels;
- The involvement of the private sector to overcome the lack of knowledge/resources/speed in public education institutions;
- The creation of motivation for public education institutions to develop more ICT-based learning services, which are more efficient and user-friendly;
- ICT-based working methods face the challenge of ensuring personal effort and genuine work products, which serve learning, for each student;
- Maintenance and renewal of ICT in educational and public institutions.

The second paper, from the above list, is an article by Ciuzas (2011), which focus mainly on assessment in conventional education. More specifically, the article aims to answer the following questions:

- Why is assessment for learning important for learning in the modern educational process?
- What are the attitudes of the teachers of Lithuanian comprehensive schools to apply assessment for learning in educational process?

The problem is addressed in three parts of the article: the first part analyses the concepts of assessment of learning and assessment for learning and reveals the importance of assessment for learning; the second part presents the methodology of the research on teachers' attitudes when assessing students' achievement to apply assessment for learning; the third part presents the results of performed research, which are:

1. Both assessment of learning and assessment for learning are important in the educational process. Assessment of learning can be used to identify the level of students' achievements by comparing it to the settled standard; however, assessment for learning is a much more powerful assessment tool if it is intended to encourage pupil's learning, improve their selfconfidence.
2. In their practice, teachers organize assessment of students' achievements themselves more frequently, rather than allow students to choose ways and forms of assessment.
3. Teachers understand the importance of assessment for learning and do not have negative attitudes related to the change of assessment process; however, the concept

of assessment for learning is not yet the practice of teachers of Lithuanian comprehensive schools.

The third paper from the above list, written by Vitienė and Mičiulienė (2008), is an article about how may a qualitative responsive information, in other words feedback, help study foreign languages more effectively when choosing eLearning course.

The article proposes a hypothesis that the clarity and precision of qualitative feedback, a systematic and individual presentation of the information as well as evaluation criteria, introduced to the students of eCourse in advance is an effective means to improve their foreign language skills.

Quasi-experiment conducted in Alytus College in 2008 revealed that the criteria defined and discussed in advance and well-timed and disciplined provision of feedback helped the majority of students improve their achievements in French continuous oral expression. Research results have validated the hypothesis that clarity and precision in qualitative feedback criteria as well as systematic and individual provision of the information is an effective means to improve students' foreign language expression skills and to form the abilities of continuous verbal expression. It is important for students to receive feedback information individually.

More specifically, **feedback** is of good quality and effective in case it:

- is focused on a particular task properly and on a regular basis;
- helps to persuade students that they have chosen the right way, induces them to correct mistakes and fulfil tasks better;
- is useful when providing students with particular help he/she needs to develop skills;
- provides a possibility for a student to foresee alternative solutions and is not limited by repeated teacher's explanations;
- helps a student acquire necessary competence.

Consequently, individual feedback presented in a qualitative way is an effective means that helps to reach for an adequate level of communication competence in the French language and simultaneously meet the students need for the most intense development of foreign language expression skills. Based on these results, feedback is of good quality and effective in case it:

- is focused on a particular task properly and on a regular basis;

- helps to persuade students that they have chosen the right way, induces them to correct mistakes and fulfil tasks better;
- is useful when providing students with particular help he/she needs to develop skills;
- provides a possibility for a student to foresee alternative solutions and is not limited by repeated teacher's explanations;
- helps a student acquire necessary competence.

Kaklauskas et al. (2010) present the fourth article of the above list. Based on them, all distance learning participants (students, professors, instructors, mentors, tutors and the rest) would like to know how well the students have assimilated the study materials being taught. The analysis and assessment of the knowledge students have acquired over a semester are an integral part of the independent studies process at the most advanced universities worldwide. A formal test or exam during the semester would cause needless stress for students.

To resolve this problem, the authors of this article have developed a Biometric and Intelligent Self-Assessment of Student Progress (BISASP) System. The obtained research results are comparable with the results from other similar studies. This article ends with two case studies to demonstrate practical operation of the BISASP System.

The first case study analyses the interdependencies between microtremors, stress and student marks. The second case study compares the marks assigned to students during the e-self-assessment, prior to the e-test and during the e-test. The dependence, determined in the second case study, between the student marks scored for the real examination and the marks based on their self-evaluation is statistically significant (the significance $>0.99\%$).

The original contribution of this article, compared to the research results published earlier, is as follows: the BISASP System developed by the authors is superior to the traditional self-assessment systems, due to the use of voice stress analysis and a special algorithm, which permits a more detailed analysis of the knowledge attained by a student.

Finally, the fifth paper compares the results of three research works carried out by the Rimkuvienė, Lepkova and Krutinis (2010). The paper focuses on the change of economic conditions after Lithuania's accession to the EU and economical recession.

The first research was carried out in 2003—before the EU accession. The research was based on a questionnaire. The second research was completed in 2006 – after the EU

accession. The third research was completed in 2009 – the economic recession was occurred. The results of three research works were compared on the basis of general indices of the economic level.

The main purpose of the comparison was to show the differences of students' opinion and the development of e-learning in the Civil Engineering Faculty of Vilnius Gediminas Technical University.

18 The case of Finland: A Success Story

Anna-Kaarina Mörsky-Lindquist (Noema-CMI Oy) conducted the literature review in Finland. There are two articles from the web and one PPT presentation, which describe Finland's school success due to formative assessment and innovation in assessment methods and tools in Finnish VET education system.

Specifically, Kati Lounema, Chief Technologist at Finnish National Board of Education, in her presentation "*Improving teaching, learning and assessment: evidence from Finland*" (presented at Cedefop Workshop, 26-27 April 2012 in Thessaloniki, Greece -available at: <http://events.cedefop.europa.eu/curriculum-innovation>

2012/images/stories/curriculum/presentations/Kati%20Lounema.pdf),

points out that in Finnish VET education system:

- Formative assessment is a part of learning/teaching process. Only summative assessment leads to validation and recognition of achieved learning outcomes.
- All units are assessed as soon as possible after studies/achievements
 - 1. Assessment during education and training
 - 2. No 'final exams' of the qualifications
- Vocational qualifications are assessed at three levels. In specialist and further vocational qualifications assessment is pass/fail.

Here is a summative table about innovation in assessment methods and tools:

	School-/curriculum based education	Competence-based qualifications
Methods for assessing learning	Verbal or written feedback on the progress of studies	
Methods for assessing competence	Vocational skills demonstrations: like competence tests	Competence tests: observation, interviews, surveys, group and self-assessment

Table 3. Assessment Innovation in Finnish VET education system:

Methods and Tools

Further interesting data about assessment in Finnish education system, comes from the online article in *The Atlantic*, which Anu Partanen –a Finnish journalist- published on December 29, 2011 (available at: <http://m.theatlantic.com/national/archive/2011/12/what-americans-keep-ignoring-about-finlands-school-success/250564/>).

By the title "*What Americans Keep Ignoring About Finland's School Success*", Partanen describes the phenomenon that the three countries that have best performing education systems in the world do not include the United States, but one of them is Finland.

She points out that "*Compared with the stereotype of the East Asian model -long hours of exhaustive cramming and rote memorization -Finland's success is especially intriguing because Finnish schools assign less homework and engage children in more creative play. All this has led to a continuous stream of foreign delegations making the pilgrimage to Finland to visit schools and talk with the nation's education experts, and constant coverage in the worldwide media marveling at the Finnish miracle.*"

She mentions that Finnish schools are all publicly financed –there are no private schools- whether for pre-K of a Ph.D. That public school system's teachers, are trained to assess children in classrooms using independent tests they create themselves. All children receive a report card at the end of each semester, but these reports are based on individualized grading by each teacher. Periodically, the Ministry of Education tracks national progress by testing a few sample groups across a range of different schools. There are no lists of best schools or teachers in Finland. The main driver of education policy is not competition

between teachers and between schools, but cooperation.

The article discusses Pasi Sahlberg's (who is director of the Finnish Ministry of Education's Center for International Mobility) statements, based to his new book *Finnish Lessons: What Can the World Learn from Educational Change in Finland?*

"Since the 1980s, the main driver of Finnish education policy has been the idea that every child should have exactly the same opportunity to learn, regardless of family background, income, or geographic location. Education has been seen first and foremost not as a way to produce star performers, but as an instrument to even out social inequality. In the Finnish view, as Sahlberg describes it, this means that schools should be healthy, safe environments for children. This starts with the basics. Finland offers all pupils free school meals, easy access to health care, psychological counseling, and individualized student guidance.

In fact, since academic excellence wasn't a particular priority on the Finnish to-do list, when Finland's students scored so high on the first PISA survey in 2001, many Finns thought the results must be a mistake. But subsequent PISA tests confirmed that Finland -- unlike, say, very similar countries such as Norway -- was producing academic excellence through its particular policy focus on equity".

At April 13, 2013 David J. Rosen posted his point of view about "*Why Finnish Schools Succeed*", (available at: <https://community.lincs.ed.gov/discussion/why-finnish-schools-succeed>) based to the article of the The Atlantic, we have just mentioned above. At his article he adds another excerpt of the leading Finnish authority on education reform, P. Sahlberg, who said in 2011:

"Americans are consistently obsessed with certain questions: How can you keep track of students' performance if you don't test them constantly? How can you improve teaching if you have no accountability for bad teachers or merit pay for good teachers? How do you foster competition and engage the private sector? How do you provide school choice?... For starters, Finland has no standardized tests. The only exception is what's called the National Matriculation Exam, which everyone takes at the end of a voluntary upper-secondary school,

roughly the equivalent of American high school. Instead, the public school system's teachers are trained to assess children in classrooms using independent tests they create themselves. All children receive a report card at the end of each semester, but these reports are based on individualized grading by each teacher. Periodically, the Ministry of Education tracks national progress by testing a few sample groups across a range of different schools."

Rosen ends up by the wish that we can all learn from each other how adult education teachers are using formative assessment, specifically and by examples, since he haven't seen many examples about that since then.

19 A global view: Application of formative assessment to eLearning

Many studies have obtained positive results with the application of formative assessment to eLearning environments (Justham & Timmons 2005; Brewer 2004; Henly 2003; Gardner *et al.* 2002; Peat & Franklin 2002; Velan *et al.* 2002; Khan *et al.* 2001).

Some researches have suggested that the development and evaluation of web-based applications should take into account different cognitive styles (Ghinea & Chen 2003; Chen & Macredie 2004; Chen *et al.* 2005). Chen and Macredie (*ibid.*) observed that web-based instructional programs are used by a number of learners who have different preferences, skills, and needs, which results in new challenges for instructional design. Zoe and DiMartino (2000) argued that further investigation into how diverse populations are using web-based instructional programs is necessary. Hence, in addition to developing a web-based formative assessment strategy, this research also explored its effectiveness in both field dependent and field independent individuals.

20 Pedagogic approaches of eLearning: The big gap in the literature review

For the most part, case studies of specific e-training programmes are descriptive, rather

than analytic or predictive, mainly located in a Higher Education, rather than vocational training environment and focused on the 'virtual classroom' model (JISC, 2007). They also tend to be restricted to particular subject areas, in particular IT, languages and engineering disciplines.

There is also an abundance of literature detailing tools for the assessment of eLearning. However, these are mainly divided into two types. Firstly there are many on-line data gathering instruments for assessing, typically, the user interface characteristics of software (e.g. student perception questionnaires) or secondly, there are devices to record and analyze usage by duration and frequency of log-in, pages accessed, user profile etc. Many of these are sophisticated in their design and ingenuity but lack guidance on interpretation and analysis.

21 Benchmarking models

According to Attwell (2006, p. 10) there have been several attempts to generate sets of criteria for quality assuring eLearning. However, these tend to be skewed towards proposing quality standards for eLearning systems and software which often disregard key variables in the wider learning environment or are based on criteria associated with evaluating traditional learning processes (and which disregard the technology) or criteria associated with measuring learner achievement through traditional pedagogies.

An additional problem is that the designers of these benchmarking systems are often locked in to a particular model of eLearning which limits their transferability.

22 Restrictions and Implementation of Assessment in the ELearning Situation

When examining the didactical aspects treated in this report, the following problematic areas for the eLearning situation can be identified:

According to Mödritscher et al. (2006) all kind of competencies – knowledge, skills, and attitudes – may be mediated within an eLearning environment. Therefore, it is possible to create learning content including facts relevant for a learner, instructions how to achieve a skill, or information about an expected behaviour. Thus, technology can be seen as an enabler for these types of competencies, because information can be enriched with multimedia assets

(Gunawardena & Mclsaac, 2003), practicing skills can be supported by using interactive elements or tutoring systems, and the behaviour of a student can be observed within the context of the eLearning system by terms of the micro-adaptive approach for eLearning (Park & Lee, 2003). In fact, it is easier to mediate knowledge through eLearning environments, while the effort for teaching skills or attitudes is much higher.

Within an eLearning system, objectives need to be defined regarding the target group. However, it is hardly possible to reach high-level learning objectives for all three types of competencies within a pure eLearning situation as stated in the study later on.

Learning objectives, which are defined by a teacher, always have to be evaluated in some way – to grade the students and to improve the quality of the course for future sessions. Considering the possibilities of eLearning, it is well documented that we can assess the gained knowledge by using limited-choice questions like quizzes or multiple-choice questions. Nevertheless, for most areas and, in particular, to reach high-level learning objectives it is necessary to examine students asking open-ended questions. Furthermore, the answers to such questions have to be interpreted and evaluated by experts. In terms of skills, we cannot measure the learning results using technology-based methods without hard efforts.

It has to be outlined that the assessment of high-level objectives can be realised in many different ways. With respect to the assessment methods focusing on didactical aspects such as defining competencies and evaluating the learning process according to the determined learning objectives, the following possibilities for implementing assessment in the eLearning situation can be found in the literature (Mödritscher et al., 2006):

First of all, most eLearning systems offer the possibilities to create and provide limited-choice questions. Although quizzes can save a lot of time to grade a large amount of students and reports about good results for low-level objectives of the cognitive domain, show a worse performance for the employment of deeper learning strategies and higher levels of cognitive processing.

Therefore, it is necessary to implement open-ended questions within the eLearning situation, for instance by tasks like writing essays or submitting some sort of project work. It is obvious that the evaluation of such tasks is extremely time-consuming for a teacher. Therefore, it is recommended to apply supporting methods such as automated grading. As an

extension of automated essay grading, Intelligent Tutoring Systems (ITS) may provide some kind of expertise within a domain and allow fully automated teaching and assessment. Yet, this kind of systems is hard to realise, often restricted to a certain domain and, thus, to a few learning objectives.

Lennon & Maurer (2003) describe several approaches beginning with the usage of professional authoring software up to a shift to the constructivistic learning paradigm. On the one side, automatically generated crossword puzzles may be enabler for the students' interest and motivation and have positive effects on assessing low to medium level objectives of the cognitive domain. On the other side, applying constructivistic learning methods is requiring a high level of students' self-motivation, but can reach high-level objectives in all domains.

For instance, one aspect of constructivism deals with collaborative learning. In particular, group activities requiring students to discuss a topic are a powerful element to extend the possibilities of eLearning. Students may treat open-ended questions, when they are working in groups.


Another interesting concept of constructivism is the so-called peer assessment. Peer assessment may reach high-level objectives for all possible domains and provide other advantages, such as using natural language processing, lowering the effort for the teacher, etc.

Finally, reports (Gredler, 2003) about games and simulation in the eLearning situation which can also be seen as a solution to reach high-level objectives, in particular for intellectual skills, but also for mediating knowledge or internalising value systems.

PART III
FINAL REMARKS

23 A Conceptual Classification of Assessment

This report has offered a presentation of assessment methods and practices across Europe, targeted specifically to those used in eLearning environments that support distance or hybrid learning. In this final chapter we define a classification framework, taking into account the purpose of the assessment and the needs of the different levels of education (i.e. high schools, universities, VET organizations), according to the literature review.

The following Table 4 contains the main assessment concepts and the level of education that according to the literature review are most commonly used. We also provide a rationalization about each assessment concept. By a symbol of  we emphasize the level of education that according to literature review the specific assessment concept is more suitable for.

ASSESSMENT CONCEPTS	LEVEL OF EDUCATION					RATIONALIZATION
	PRIMARY	SECONDARY	HIGHER	ADULT	VET	
MULTIPLE CHOICE QUESTION (MCQ)	✓	✓		✓		MCQ are applied to reach lower-level objectives, like recalling facts
TRUE/FALSE (T/F)	✓	✓		✓		T/F items seem to be a measure of more basic knowledge
SHORT ANSWERS	✓	✓	✓	✓	✓	Require students to formulate their own answers, which do not have to be pre-determined
E-PORTFOLIO	✓ ★	✓	✓	✓	✓	★ Plus access for parents in primary and secondary education
TABLES & CHARTS	✓	✓				Provide better visualization

Table 4. A Conceptual Classification of Assessment

ASSESSMENT CONCEPTS	LEVEL OF EDUCATION					RATIONALIZATION
	PRIMARY	SECONDARY	HIGHER	ADULT	VET	
PEER - ASS/NT		✓	✓	✓	✓	Encourage learners to take responsibility for their own learning, and deep rather than surface learning
DIAGNO- STIC ASS/NT	✓ [★]	✓	✓	✓	✓	[★] Plus early identification of learning difficulties
FORMATIVE ASS/NT Or ASS/NT FOR LEARNING	✓	✓	✓	✓	✓	Assessment works best when it is on going, not episodic
	LEVEL OF EDUCATION					

ASSESSMENT CONCEPTS						RATIONALIZATION
	PRIMARY	SECONDARY	HIGHER	ADULT	VET	
SUMMATIVE ASS/NT Or ASS/NT OF LEARNING			✓	✓	✓	Usually leads to formal qualification or certification of a skill
OBSERVA- TION IN FORA, WIKIS, WEBLOGS	✓	✓	✓	✓	✓	eLearning requires students to interact with other students in the chat room
GROUP WORK		✓ ✘				✘ Secondary education features a collaborative approach to learning, with a strong emphasis on analysis and discussion


ASSESSMENT CONCEPTS	LEVEL OF EDUCATION					RATIONALIZATION
	PRIMARY	SECONDARY	HIGHER	ADULT	VET	
AUTHENTIC LEARNING	✓	✓	✓	✓	✓	The learning task should be based on real life problem solving with a meaningful context for planned learning experiences
MULTIPLE ASS/NT METHODS	✓	✓	✓	✓	✓	The use of multiple methods of assessment can enhance reliability
FEEDBACK	✓	✓	✓	✓ 	✓	Comments on learner's work that show them how to move on to the next stage of achievement.

Table 4. A Conceptual Classification of Assessment

References

- Angelo, T.A. & K.P. Cross (1993). *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco: Jossey-Bass Publishers.
- Barron, B.J.S., Schwartz, D.L., Vye, N.J., Moore, A., Petrosino, A., Zech, L. & Bransford, J. (1998). Doing with understanding: lessons from research on problem- and project-based learning. *Journal of the Learning Sciences*, 7, pp. 271–311.
- Beale, S. (2007). *Blogs, reflective practice and student-centered learning*, Advanced Interaction Group School of Computer Science, University of Birmingham. Retrieved April 29, from http://www.bcs.org/upload/pdf/ewic_hc07_sppaper1.pdf
- Bell, B. & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85, pp. 536–553.
- Bennett, R. E. (1993). On the Meaning of Constructed Response. In R. E. Bennett, Ward, W. C. (Ed.), *Construction versus Choice in Cognitive Measurement: Issues in Constructed Response, Performance Testing, and Portfolio Assessment* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Biggs, J. (1999). *Teaching for quality learning at University*. Buckingham & Philadelphia: SRHE & Open University.
- Bransford J.D., Brown A. & Cocking R. (2000) *How People Learn: Mind, Brain, Experience and School, Expanded Edition*. National Academy Press, Washington, DC.
- Brewer, C.A. (2004). Near real-time assessment of student learning and understanding in Biology courses. *Bioscience*, 54, pp. 1034–1039.
- Brown, S. & Knight, P. (1994). *Assessing Learners in Higher Education*. Kogan Page, London.
- Chen, S.Y. & Macredie, R.D. (2004). Cognitive modelling of student learning in web-based instructional programmes. *International Journal of Human–Computer Interaction*, 17, pp. 375–402.
- Chen, S.Y., Magoulas, G.D. & Dimakopoulos, D. (2005). Aflexible interface design for web directories to accommodate different cognitive styles. *Journal of the American Society for Information Science and Technology*, 56, pp. 70–83.
- Cronin, J.C., (1993). Four misconceptions about authentic learning. *Educational*

Leadership, 50(70), 78-80.

- Davi, A., Frydenberg, M. & Gulati, G.J. (2007). Blogging Across the Disciplines: Integrating Technology to Enhance Liberal Learning. *MERLOT Journal of Online Learning and Teaching 3(3)*.
- Drexler, W., Dawson, K. & Ferdig, R.E. (2006). Collaborative blogging as a means to develop elementary expository writing skills. *Electronic Journal for Integration of Technology in Education, 6*, pp. 140-150.
- Ebel, R.L. & Frisbie, D.A. (1991). *Essentials of Educational Measurement*, 5th edn. Prentice Hall, Inc, Englewood Cliffs, NJ.
- Embretson, S. (2002). Generating Abstract Reasoning Items with Cognitive Theory. In S. Irvine, Kyllonen, P. (Ed.), *Item Generation for Test Development* (pp. 219–250). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Ghinea, G. & Chen, S.Y.(2003). The impact of cognitive styles on perceptual distributed multimedia quality. *British Journal of Educational Technology, 34*, pp. 393–406.
- Glaser, R. (1991). Expertise and Assessment. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and Cognition*. Englewood Cliffs, NJ: Prentice-Hall.
- Glaser, R. (1988). Cognitive and environmental perspectives on assessing achievement. Paper presented at the *Assessment in the Service of Learning ETS Invitational Conference*, Princeton, NJ.
- Gredler, M.E. (2003). Games and simulations and their relationship to learning. In: *Educational Technology Research and Development*, pp. 571-582.
- Gronlund, N. E. (2003). *Assessment of Student Achievement*. New York: Pearson Education, Inc.
- Gruttmann, S. & Usener, C. (2011). *Prüfen mit Computer und Internet - Didaktik, Methodik und Organisation von E-Assessment*. University of Graz
- Gunawardena, C.N. & Mclsaac, M.S. (2003). Distance education. In: *Educational Technology Research and Development*, pp. 355-396.
- Haynes, M., Anagnostopoulou, K., et al. (2004). *eLearning – A tutor guide*. Middlesex University Press.
- Henly, D.C. (2003). Use of Web-based formative assessment to support student learning in a metabolism/nutrition unit. *European Journal of Dental Education*,

7, pp. 116–122.

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), pp. 23-48.

IDS (2002). *Instruction at FSU: A Guide to Teaching and Learning Practices*. Instructional Development Services, Florida State University.

Jain, L.C., Howlett, R.J., Ischalkaranje, N.S., and Tonfoni, G. (2002). Virtual Environments for Teaching & Learning. In: *Series of Innovative Intelligence*, Vol. 1

Justham, D. & Timmons, S. (2005). An evaluation of using a web-based statistics test to teach statistics to post- registration nursing students. *Nurse Education Today*, 25, pp. 156–163.

Khan, K.S., Davies, D.A. & Gupta, J.K. (2001). Formative self- assessment using multiple true-false questions on the Inter- net: feedback according to confidence about correct knowledge. *Medical Teacher*, 23, pp. 158–163.

Khom, W. (2006). *E-Testing. Die konsequente Fortführung von e-Learning*. bit Media.

Lennon, J. & Maurer, H. (2003). Why it is Difficult to Introduce eLearning into Schools And Some New Solutions. In: *Journal of Universal Computer Science (J.UCS)*, Vol. 9, pp. 1244-1257.

Lombardi, M. (2007). *Authentic Learning for the 21st Century: An Overview*. ELI Paper. EDUCAUSE.

Mabrito, M. (2004). Guidelines for establishing interactivity in online courses. *Innovate*, 1(2). Retrieved May 15, from <http://www.innovateonline.info/index.php?view=article&id=12>

McDonald, M. E. (2002). Developing Multiple-Choice Items. In *Systematic Assessment of Learning Outcomes* (pp. 83–120). Sudbury, MA: Jones and Bartlett Publishers.

McMahon, M. (2010). Ponderers, Sloggers, Slackers and More: Understanding the profiles of student bloggers to help promote academic self-regulation. In C.H.

- Steel, M.J. Keppell, P. Gerbic & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future. Proceedings ascilite Sydney 2010* (pp.620-629). Retrieved April 20, from <http://ascilite.org.au/conferences/sydney10/procs/Mcmahon-full.pdf>
- McLoughlin, C. (2002). Learner Support in Distance and Networked Learning Environments: Ten Dimensions for Successful Design. *Distance Education*, 23(2), pp. 149-162.
- Mislevy, R. J. (1996). Test Theory Reconceived. *Journal of Educational Measurement*, 33, pp. 379–417.
- Mödritscher, F., García-Barrios, V.M., & Gütl, C. (2004). The Past, the Present and the Future of adaptive E- Learning: An Approach within the Scope of the Research Project AdeLE. In: *Proceedings of the International Conference on Interactive Computer Aided Learning (ICL2004)*, Villach, Austria.
- Mödritscher, F. & Sindler, A. (2005). Quizzes are not enough to reach high-level learning objectives. In: *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2005)*, Montreal, Canada, pp. 3275-3278.
- Nunes, M. & McPherson, M. (2006). Learning support in online constructivist environments in information systems. *HEA-ICS Electronic Journal*, 5(2). 1-11.
- Osterlind, S. J. (1998). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats*. Norwell, MA: Kluwer Academic Publisher.
- Park, O. & Lee, J. (2003). Adaptive Instructional Systems. In: *Educational Technology Research and Development*, pp. 651-684.
- Parshall, C. G. (2002). Item Development and Pretesting in a CBT Environment. In C. Mills, Potenza, M., Fremer, J., Ward, W. (Ed.), *Computer-Based Testing* (pp. 119–141). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Parshall, C. G., Davey, T., & Pashley, P. J. (2000). Innovative Item Types for Computerized Testing. In W. Van der Linden, Glas, C. A. W. (Ed.), *Computerized Adaptive Testing: Theory and Practice* (pp. 129–148). Norwell, MA: Kluwer Academic Publisher.
- Parshall, C. G., Spray, J., Kalohn, J., & Davey, T. (2002). Issues in Innovative Item Types. In *Practical Considerations in Computer-Based Testing* (pp. 70–91). New

York: Springer.

Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). Knowing What Students Know: The Science and Design of Educational Assessment. In *N. R. C. Center for Education* (Ed.). Washington, D.C.: National Academy Press.

Perrenoud, P. (1998) From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in Education*, 5, pp. 85–102.

Raith, A. (2004). *E-Testing im österreichischen Hochschulbereich – eine Chance?* Donau-Universität Krems.

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 Assessments: Alternative Views of Aptitude, Achievement and Instruction* (pp. 37–76). Boston, MA: Kluwer Academic Publishers.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, pp. 119–144.

Scalise, K. & Gifford, B. (2006). Computer-Based Assessment in E-Learning: A Framework for Constructing “Intermediate Constraint” Questions and Tasks for Technology Platforms. *Journal of Technology, Learning, and Assessment*, 4 (6).

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, Vol. 35, pp. 453-472.

Shepard, L. (1991a). Interview on assessment issues with Lorrie Shepard. *Educational Researcher*, 20(2), 21–23, 27.

Shepard, L. (1991b). Psychometricians' beliefs about learning. *Educational Researcher*, 20(6), 2–16.

Spector, J.M. & Ohrazda, C. (2003). Automating instructional design: Approaches and limitations. In: *Educational Technology Research and Development*, pp. 685-700.

Stout, W. (2002). Test Models for Traditional and Complex CBTs. In C. Mills, Potenza, M., Fremer, J., Ward, W. (Ed.), *Computer-Based Testing* (pp. 103–118).

Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Williams, C.E. (2004) School district of Philadelphia uses Web-based system to increase student achievement. *T H E Journal* 31, pp. 51–52.

Williamson, I. J. Bejar & R. J. Mislevy (Eds.), *Automated Scoring of Complex Tasks in Computer Based Testing*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Zoe, L.R. & DiMartino, D. (2000). Cultural diversity and end- user searching: an analysis by gender and language background. *Research Strategies*, 17, pp. 291–305.

ANNEX

More evidence about E-assessment: The case of UK in numbers

JISC (2007, p. 11) publication reports some key steps towards e-assessment in England.

Based to those facts:

- Over 80,000 e-assessments were performed at Loughborough University during 2006.
- The City & Guilds Institute has delivered 1.3 million tests since 2003, 650,000 of which were in the academic year 2005-06.
- The CCEA GCE Moving Image Arts completes its pilot phase in 2007-08 to become the first AS/A2 qualification to be assessed entirely online.
- Awarding bodies in England offer an online GCSE science in 2006.
- In 2006, awarding bodies and TechDis collaborate on the production of guidelines for accessibility in e-assessment.
- A pilot online multiple-choice examination in biotechnology at Intermediate 2 and Higher Level was offered by SQA for the first time in 2006.
- In 2007, the Phase 2 report became available from Project e-Scape on e-portfolio-based assessment of GCSE design and technology coursework.
- SQA now provides online National Qualification Unit assessments in computing and higher mathematics. Pupils with additional assessment requirements at Standard Grade are able to word process their responses.
- Following initial trials in 2001, on-demand basic and key skills tests in England, Wales and Northern Ireland can be taken online at levels 1 and 2.
- Availability of VLEs and their use by UK educational providers grows year on year.

GLOSSARY

Assessment	Defines what is taught and how it is learnt.
Authentic Assessment	Assessment where students engage in learning that has real world relevance
Completion question	Question type which involves items in which not all the components for the answer are supplied in advance, or in which the entire problem-space is contained in the set of answers
Criteria for success	The minimal level of mastering for each learning objective in percentages.
Diagnostic assessment	Assessment of a learner's knowledge and skills at the outset of a course.
e-Portfolio	An e-portfolio is a deliberate choice of electronic documents and certificates.
Formative assessment	Assessment that provides developmental feedback to a learner on his or her current understanding and skills.
High-Value Assessment	Assessment which coalesce plenty of complex learning.
Learning objective	Describes the level of knowledge, skills, competences and other characteristics that any learner should achieve in the process of learning.
Low-Value Assessment	Assessment which is low-risk for the learner and which does not involve a lot of points (or no points, or only extra-credit points).
Matching type question	A question which requires matching one term, phrase or image to another by the student.
Multiple Choice Questions (MCQ):	A question with several options that are considered "correct" and students may choose more than one.
Peer Assessments	Shared learning spaces where peers may post and share their work.
Presentation/portfolio	A wide variety of complex performances that include such activities as projects, portfolios, demonstrations, experiments, fine art performances, and medical diagnoses or other professional practicum as well as teaching and extended group activities, discussions, and interviews.
Summative assessment	The final assessment of a learner's achievement, usually leading to a formal qualification or certification of a skill.

True/False (TF) questions

Questions ask the student to determine if a statement is true or false.

Wikis

Wikis are asynchronous, editable web pages that enable collective text-writing, the compilation of knowledge and project management.