

## **PROMOTING E-ASSESSMENT QUALITY IN HIGHER EDUCATION: A CASE STUDY IN ONLINE PROFESSIONAL DEVELOPMENT**

Luís Tinoca  
University of Lisbon  
Portugal

### **Abstract**

The main purpose of this paper is to present and discuss a conceptual framework for e-assessment intended to enhance the quality and facilitate the design of a Competence Assessment Program (CAP) (Baartman, Bastiaens, Kirschner, & Vleuten, 2007) for online virtual environments. The main intended outcome is the improvement of the presented e-assessment framework through the feedback received from the participants, and the enhancement of the proposed CAP for the next iterations of the design-based research method being used.

### **The Challenge of E-Assessment**

Universidade Aberta, the Portuguese Open University, adopted in 2007 a fully virtual pedagogical model for online education (Pereira, Mendes, Morgado, Amante, & Bidarra, 2007), aligned with the most recent developments in education theory with a particular focus on e-learning. Simultaneously, in Europe, the Bologna Process (European Commission, 2008), challenged higher education to promote learning environments that are centered in the development of competences. This challenge requires not only a shift in the perceived goals of higher education, but also in the selection of the methodologies to be used, as well as a drastic change in the assessment strategies to implement in order to foster a competence based curriculum. Within this setting, at Universidade Aberta, we were confronted with the need to reflect on this issue, taking into account the latest developments in the field of competence assessment with an edumetric framework, particularly in the case of e-assessment in online environments.

To answer this challenge we have developed a new conceptual framework for e-assessment, composed of four dimensions identified as especially relevant for online contexts, and further operationalized into fifteen criteria (see figure 1). The four proposed dimensions for e-assessment to be considered in the definition of competence based e-assessment strategies are: authenticity, consistency, transparency, and practicability. The authenticity domain emphasizes the need to warrant that online assessment tasks are complex, related to real life context, and recognized as significant by students, teachers and employers. Consistency stresses the importance of aligning the

competences being assessed with the e-assessment strategies being used and the assessment criteria, as well as the need to use a variety of indicators. The transparency dimension promotes student engagement in online tasks through the democratization and visibility of the e-assessment strategies being used. Finally, practicability is particularly important in online contexts given their specificities considering resources, time and training costs, as well as their efficiency and sustainability.

It is also important to notice that these dimensions are articulated, representing several degrees of reciprocal interdependence. The dimension of practicability, for example, frequently neglected, may have a decisive influence in the level of implementation of the remaining dimensions. The criteria included in each dimension are important not only as contributors to the characterization of each of the dimensions, but also to illustrate their degree of implementation.

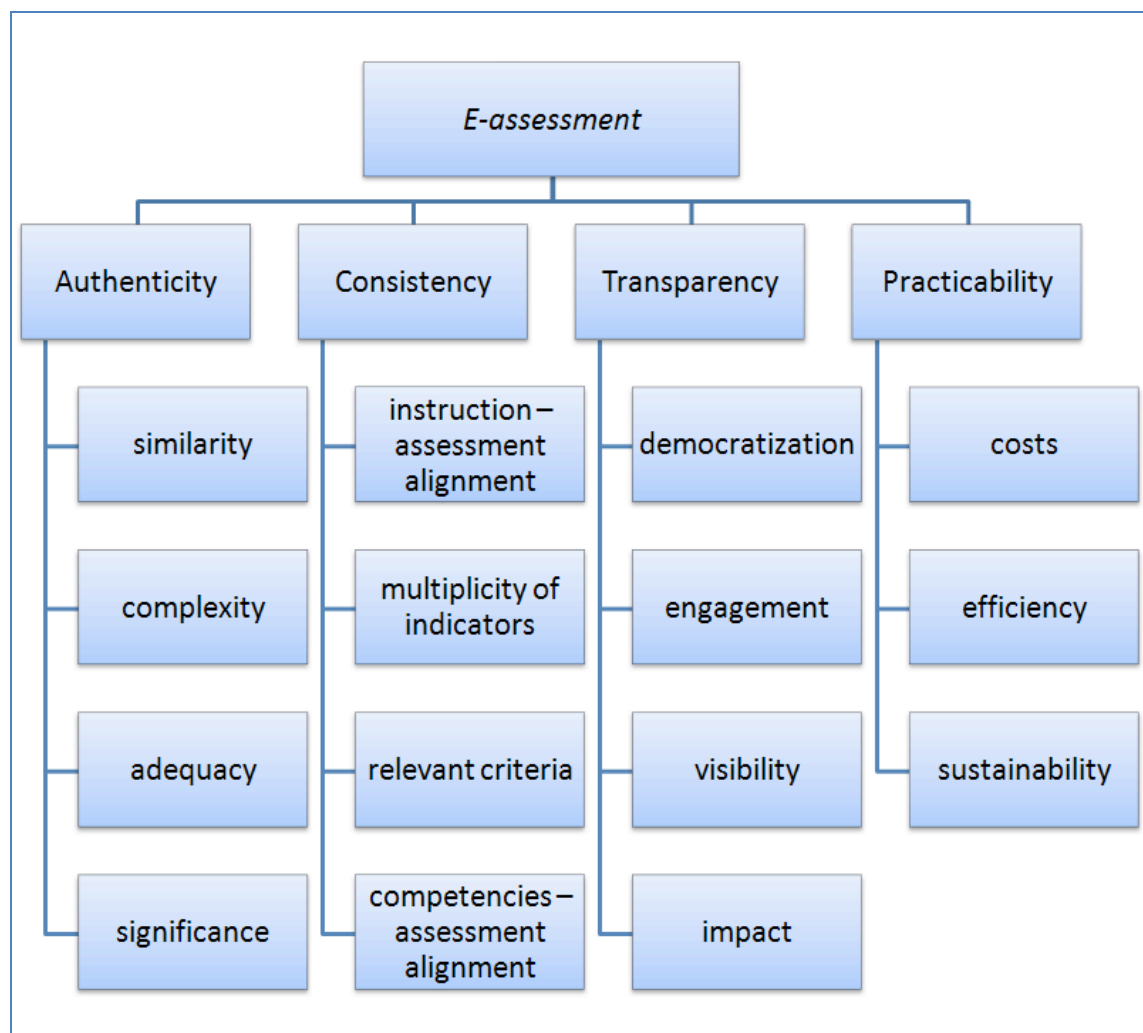


Figure 1. A conceptual framework for e-assessment.

The concept of *authenticity* is related to the degree of similarity between the competences being assessed and the ones required in real/professional life. In this dimension are included four reference criteria contributing to the degree of assessment authenticity:

*Similarity* – refers to the way in which the assessment strategy is related to the real life context (physical and social), meaning that assessment should reflect the competences needed in real/professional life (Baartman et al., 2007, Dierick and Dochy, 2001; Gulikers et al., 2004). The physical context refers to the type and number of available digital resources, while the social context is supposed to be aligned with the equivalent social processes in a real/professional situation.

*Complexity* – refers to the nature of the e-assessment tasks, more specifically, to the cognitive challenges that are imposed by its resolution/development, similar to real/professional, daily life challenges often ill-structured and with a variety of possible solutions (Gulikers et al., 2004; Herrington & Herrington, 1998; Mateo & Sangrà, 2007).

*Adequacy* – is connected with the need to provide adequate performing conditions (time, digital resources, etc.) for the e-assessment tasks, in accordance with their complexity. This often implies the removal of unrealistic restraints imposed by formal educational contexts (Gulikers et al., 2004; Herrington & Herrington, 1998). The participants' responses should explain understandings and conclusions; they should be clear, coherent and provide richness in details, qualifications and argument (MacLellan, 2004b).

*Significance* – includes the significant value of the e-assessment task for students, instructors and employers (Baartman et al., 2007; Gulikers et al., 2004). McDowell (1995) considers that the connection between the e-assessment tasks and the learning needs should be clear and perceived by the students/learners.

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

*Instruction-assessment alignment* – refers to the need to provide e-assessment scenarios that are representative of the learning situations experienced by the students (Palm, 2008), warranting in this way the agreement between the work developed during the instruction and the assessment tasks being used.

*Multiplicity of indicators* – this criterion is related to the need of employing a variety of e-assessment methods, contexts, moments, and assessors (Dierick & Dochy, 2001; Herrington & Herrington, 1998). This way, the usage of self-assessment, peer-assessment, besides instructor lead assessment, as well as a diversity of methods and tasks, should be supported in a variety of moments during the learning process. This approach further contributes to the competence assessment program equity by allowing for diverse means of participation.

*Relevant criteria* – considers the relevance of the assessment criteria being used for the competences that are being assessed (individual or collaborative) (Herrington & Herrington, 1998; Pereira, Tinoca, & Oliveira, 2009).

*Competences-assessment alignment* – is related to the need of assuring the coherency between the competences that are intended to be developed and the assessment design being used (Palm, 2008; Pereira, Tinoca, & Oliveira, 2009).

The *transparency* dimension intends to make the entire competence assessment program visible and comprehensible for all participants. For this reason, it is important that the students/learners are able to understand the fairness of their assessment, as precisely as their instructors, requiring them to have a complete knowledge of all assessment criteria and their relative weights. According to Dierick and Dochy (2001) and McDowell (1995), the clarification of the assessment criteria coupled with the fact that the students/learners have prior knowledge of the competence assessment program has a positive impact on their learning. For this dimension four criteria were considered:

*Democratization* – refers to the availability and possible participation in the definition of the assessment criteria (Dierick & Dochy, 2001). Moreover, the students/learners should know, from the onset, what the assessment goals are and who their assessors are going to be. This way, they know immediately what is expected from them, and can adjust their learning processes accordingly (McConnell, 2006).

*Engagement* – is related to the availability and possible participation in the definition of the learning goals (MacLellan, 2004) and performance criteria (Pereira, Tinoca, & Oliveira, 2009). It allows students/learners to participate in the definition of their learning environment, further contributing to their active participation, commitment and responsibility (Pereira, Tinoca, & Oliveira, 2009).

*Visibility* – refers to the possibility of presenting/sharing their learning processes and/or products with others (peers, assessors, instructors, community, etc.) (Gulikers et al., 2004).

*Impact* – is related to the effects that the e-assessment strategies have in the learning process and in the design of the educational program (Baartman et al., 2007). According to Brinke (2008) the assessment design must have a positive impact in the learning process.

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

*Costs* – referring to the time costs (for both assessors and students/learners) (Brinke, 2008; Dierick & Dochy, 2001), as well as the digital resources or additional investments, particularly training, that may be necessary to implement the assessment design.

*Efficiency* – goes beyond costs, to consider the relationship costs/effects of the e-assessment design for the institutions, instructors and learners, particularly considering the expected results. It should, for example, be translated into the choice of certain technological tools aimed at promoting a more efficient assessment (Brown, 2004).

*Sustainability* – is related to the need of assuring that it is possible to successfully implement and sustain the proposed assessment design, taking into account the learner profiles (level of education, previous training, familiarity with the assessment tools, prior knowledge, competences, etc.) and the contextual constraints, both for the organizations and for the assessors.

The presented dimensions, and criteria, can be used, in this way, as a framework for the development of a competence assessment program in an online context. What is more, they serve as guidelines for the quality evaluation of the e-assessment strategies being used. The discussed criteria, more than just illustrating the different features of each dimension, allow for the operational description of each criterion stage of implementation, and so contribute to the evaluation of the achieved assessment strategy quality level.

The suggested dimensions and criteria cover not only the six aspects of construct validity proposed by Messick (1994, 1995), but also the ten quality criteria for competence assessment programs suggested by Baartman et al. (2007). However the proposed framework emphasizes the importance of four dimensions especially relevant for online contexts in general and at the Universidade Aberta in particular. Furthermore, the fifteen suggested criteria operationalizing these dimension more than subsuming the criteria from previous frameworks, advance new aspects particularly relevant for e-assessment such as the adequacy to online contexts, the distinction between instruction-assessment alignment and competences-assessment alignment, democratization, engagement, visibility and sustainability.

### **The Context**

The *Information and Communication Technologies Practice* (ICTP) course is part of the graduate program on teacher training and professionalization at the Portuguese Open University. This program enrolls k-12 teachers, with a minimum of 5 years of experience but without any pedagogical training. The goal of this program is to provide these teachers with the pedagogical competences at they are lacking. The ICTP course was created within this context as a general educational course intended to provide the teachers with the necessary competences to use ICT in their classroom practices.

### **The Pedagogical Model**

All students at the Portuguese Open University are enrolled in fully online programs, created under the pedagogical model for online teaching (Pereira et

al., 2007) designed and approved in 2007. This model is established as a framework for the organization and development of all programs and courses at this university, including two variations: one for undergraduate courses with large numbers of students per class; and one for graduate courses with a small number of students per class (a maximum of 30).

In the case of this teacher training program it was opted for a hybrid version of this model, given its graduate nature on the one hand, and the fact that there were a very high number of participants (635) as a result of the protocol that was established with the Portuguese Ministry of Education and the teachers union.

This pedagogical model is based on four cornerstones:

- Flexibility
- Student-centered learning
- Interaction
- Digital inclusion

In particular, the flexibility cornerstone advocates for an asynchronous mode of interaction, with well-established moments for communication between student and with the instructors throughout the semester. The student-centered learning cornerstone is implemented through a series of proposed activities for the students to complete during the semester, supported by a variety of online pedagogical resources. Also, to facilitate the students organization they are provided with a Course Plan (CP) at the beginning of the semester, where all competences to be developed are detailed, as well as the course contents, the assessment activities, and a weekly schedule. Throughout the semester the students are then given access to several sets of formative activities accompanied by the instructor's suggestions to facilitate their completion. In this way, the students are proposed a learning path, guided and scheduled from the beginning by the CP, implying a continuous learning process, both independently and collaboratively with their colleagues, marked by the completion of the formative activities and online interaction forums both the choice is made by the students during the initial three weeks of each semester using an online tool available in each course.

The digital inclusion cornerstone, implicit in this model where all activities are completed online with the exception of the final exam, also requires that the students, before initiating any program at the Portuguese Open University must attend and complete a previous module entitled *Online Familiarization*, lasting two weeks, and where they develop competences particularly relevant for online interaction and study.

The discussed pedagogical model calls for two main types of instruments to be used in undergraduate fully online courses: e-folios and p-folios. The e-folio "is a short digital document elaborated by the student and published online to be visualized by the teacher, and should clearly demonstrate that the student acquired or developed a given competence" (Pereira, et al. 2007, p. 19). Here are some examples of what an e-folio may include: a critical reflection about

the student's own learning, a report about field work, a problem solving activity, a reading review, or the production of an artifact. Being introduced as an instrument for continuous assessment of the competences promoted in a particular curricular unit, students are expected to produce two or three e-folios per semester. The e-folios may be complemented by a p-folio that takes place in a face-to-face setting. The p-folio may take the form of "a set of questions defined by the teacher, or other forms, such as the presentation of a project or a report, according to the competences to be developed by the students" (Pereira et al. 2007, p. 19).

Two main research questions were asked: 1) "Do the participating teachers recognize de proposed CAP as a consistent, fair and relevant assessment strategy for their learning?" "And, 2) "Are the proposed digital assessment quality dimensions (authenticity, consistency, transparency, and practicability) reflected in the proposed assessment activities?"

### **Method**

The methodology used in the whole project is design-based research. Bereiter (2002) defines design-based research as any kind of research that produces findings that are fed back into further cycles of innovative design. The use of design-based research includes five main principles: strategic investigation; contextualized research focused on the identified problems; research targeted at innovation and practical applications; a practice conceived and oriented by research goals; the existence of research and assessment cycles for the interventions. Anderson (2005) specifically refers to the suitability of applying design-based research methods to educational innovation contexts, in general, and to distance education, in particular. Moreover, in order to assess the participants understanding of the CAP and of each assessment activity, two main strategies were used: a large scale survey (n=230), to be completed at the end of the semester; and focus interviews with teachers (n=5) particularly selected for their critical comments of the assessment strategies used.

### **Results**

To answer the first research question about the quality of the CAPs, the quantitative data (scores in the web-based questionnaire) were analyzed using SPSS. The internal consistency of the scales of all indicators appeared to be moderate to good (Cronbach's Alpha ranging from .61 to .88). Means and SDs for the 4 dimensions and 15 criteria were calculated. In order to assess differences between dimensions and criteria an overall mean score was calculated. One-sample T-tests were then used to compare the scores with the overall mean.

The overall results (see table 1) showed good values for the dimensions of authenticity, consistency and transparency, even though some criteria scored lower than others. In particular, the criteria adequacy, significance, and especially engagement presented lower scores. The dimension of practicability scored especially low in all its criteria.

Table 1

*Means and Standard Deviations for Each Dimension and Assessment Criteria*

n = 230 Dimensions and Criteria	Score (1-4)	
	M	SD
<b>Authenticity</b>	<b>3,1</b>	<b>0,4</b>
Similarity	3,3	0,3
Complexity	3,2	0,5
Adequacy	2,9	0,4
Significance	2,8	0,4
<b>Consistency</b>	<b>3,2</b>	<b>0,4</b>
Instruction-assessment alignment	3,1	0,6
Multiple indicators	3,1	0,4
Relevant criteria	3,3	0,3
Competences-assessment alignment	3,4	0,4
<b>Transparency</b>	<b>3,3</b>	<b>0,45</b>
Democratization	3,5	0,3
Engagement	2,4	0,5
Visibility	3,3	0,5
Impact	3,3	0,4
<b>Practicability</b>	<b>1,9</b>	<b>0,9</b>
Costs	1,8	0,9
Efficiency	2,2	1,1
Sustainability	1,9	0,8

Qualitatively, the interviews with the teachers' data analysis involved iterative analysis and revision of the coding scheme. During the analysis the initial coding key was revised to account for emergent sub-codes and all of the data was recoded using a final coding scheme. Some of the codes were quantified in order to foster a more meaningful comparison of the data by allowing patterns to be identified and further explored (Chi, 1997). The data was processed using the NVIVO qualitative data treatment software.



From this analysis four main categories emerged. First, the most representative was the teachers' recognition for the authenticity of the proposed themes and their similarity to their real/professional life contexts. Second, concerning the resources available online the overall appreciation is also very positive, being acknowledged by the participating teachers as relevant for their practice and fueling their self-assessment and professional development. Third, the e-folio tool is recognized as very closely related to the 4 proposed dimensions; on the other hand, the p-folio, even though recognized as practical and consistent is challenged for its transparency and authenticity. Finally, concerning the adopted CAP and the proposed assessment activities, even though the participants recognize them as transparent, consistent and authentic, it also clearly emerges that the practical requirements imposed by this pedagogical model – especially the absence of assessment for the online interaction forums given the high number of participants – is reflected in their less positive comments.

### Conclusion and Discussion

Given the nature of the design-based research method being used there are two main goals to fulfill. On the one hand, regarding our theoretical development of the conceptual framework for e-assessment it is now important to revisit it and reconsider the proposed dimension and criteria, their relationships and relevance. In particular, careful thought should be given to the dimension of practicability where the lower scores were attained. Is it properly defined? Should it be assessed differently? The authors' (Pereira, Oliveira & Tinoca, 2010) initial intention with this dimension was to use it mainly as a reference for the instructors developing e-assessment strategies and here it was evaluated only from the standpoint of the participating teachers, this may be one of the reasons responsible for the observed gap.

On the other hand, the attained results should prompt our reflection and revision of the proposed CAP and in particular the e-assessment strategies proposed for this course. Even though the authenticity, consistency and transparency dimensions had high scores, the results across the proposed criteria were not even. Some of them scored significantly lower than others. Particular attention should now be given to the criteria that scored at the lower ranges. The *engagement* criteria, for example, had a particularly low score when compared with the others in its dimension. From the data gathered with the interviews we hint that this may be given to the fact that the participating teachers did not recognize the “p-folio” as transparent and authentic when compared with the other practices throughout the course. This was given not only to its nature, but also to its very high weight towards the final grade. Also, despite what was previously said, the practicability criteria should also prompt our reflection over the proposed CAP. Why are the teachers considering the costs, efficiency and sustainability so poorly? What can be done to improve this perception?

However, it must be recognized that the overall appreciation of the participating teachers was quite positive as emerged from the focus interviews. They not only recognized the course as relevant and impacting their practice,

but also they considered the CAP as fair and relevant for their learning and professional development.

### References

- Anderson, T. (2005). Design-based research and its application to a call centre innovation in distance education. *Canadian Journal of Learning and Technology*, 31(2).
- Baartman, L.K.J., Bastiaens, T.J., Kirschner, P.A. & Vleuten, C. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, 2, 114-129.
- Bereiter, C. (2002). Design research for sustained innovation. *Cognitive Studies, Bulletin of the Japanese Cognitive Science Society*, 9(3), 321-327.
- Brinke, D. (2008). *Assessment of prior learning*. Maastricht, the Netherlands: Datawyse.
- Brown, S. (2004). Assessment for learning. *Learning and Teaching in Higher Education*, 1, 81-89.
- Dierick, S., & Dochy, F. J. R. C. (2001). New lines in edumetrics: New forms of assessment lead to new assessment criteria. *Studies in Educational Evaluation*, 27, 307-329.
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Design*, 53, 67-87.
- Herrington, J. & Herrington, A. (1998). Authentic assessment and multimedia: How university students respond to a model of authentic assessment. *Higher Education Research & Development*, 17(3), 305-322.
- Maclellan, E. (2004). How convincing is alternative assessment for use in higher education? *Assessment & Evaluation in Higher Education*, 29(3), 311-321.
- McConnell, D. (2006). *E-learning groups and communities*. Berkshire: Open University Press.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13-23.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment Research & Evaluation*, 13(4). Retrieved from <http://pareonline.net/getvn.asp?v=13&n=4>
- Pereira, A., Oliveira, I., & Tinoca, L. (2010). A cultura de avaliação: Que dimensões? In F. Costa, G. Miranda, J. Matos, I. Chagas & E. Cruz (Eds.). *Actas do I Encontro Internacional TIC e Educação: TICeduca*.
- Pereira, A., Oliveira, I., Tinoca, L., Amante, L., Relvas, M.J., Pinto, M.C. & Moreira, D. (2009). Evaluating continuous assessment quality in competence-based education online: The case of the e-folio. *European Journal of Open, Distance and E-learning*. EDEN.

- Pereira, A., Tinoca, L., & Oliveira, I. (2010). Authentic assessment contribution to competence based education at Universidade Aberta: Questions and challenges. In S. Mukerji and P. Tripathi (Eds.), *Cases on technological adaptability and transnational learning: Issues and challenges*. IGI Global.
- Tinoca, L., Oliveira, I., & Pereira, A. (2010, April). Online group work patterns: How to promote a successful collaboration. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the Seventh International Conference on Networked Learning*. Aalborg, Denmark.