DESIGN OF AN ADAPTIVE E-ASSESSMENT SYSTEM FOR FORMATIVE ASSESSMENT: TRAiL

Nuno Dorotea and Neuza Pedro
University of Lisbon
Portugal

Abstract

Formative assessment plays a key role in the teaching and learning processes. It informs about the status of students' learning and helps to identify difficulties, promoting students' autonomy and learning self-regulation. At the same time, it enables teachers to evaluate their students' progress. As a result of a study developed between 2014 and 2017 with Portuguese primary and secondary school teachers, this paper presents a proposal to develop a new Adaptive e-Assessment System that can support the implementation of formative tests and, consequently, contribute to the academic success of students.

Keywords: formative assessment, online digital tests, feedback, adaptive e-assessment system

Introduction

Formative assessment has proven to play a key role in enhancing the teaching and learning processes (Marzano, 2007). According to Black and Wiliam (1998), when presented as a systemic and regular practice, it improves student learning and his outcomes. Formative assessment consists of the gathering, analysis and interpretation of systematic evidence to determine how students' learning is matching teacher expectations and curricular objectives. It reports students' progress, where they face problems and how well they are progressing.

Tests, as a formative activity, allow the assessment of several curricular contents as well as the evaluation of many students simultaneously. According to Roediger and Butler (2013), the regular use of tests and the feedback provided by them are factors that enhance student learning. At the same time, relevant and just-in-time feedback promotes student confidence, autonomy, and self-regulation of his/her own learning processes (Gilbert, Whitelock, & Gale, 2011). Tests also allow teachers to reflect and make decisions about the teaching process (Earl, 2003), promoting the adjustment of strategies and methodologies according to the students' achievements and difficulties.

We can consider effective feedback as one that promotes the self-regulation of learning by students, as well as the readjustment of the teaching process by teachers. According to Irving (2007), to be effective, feedback must be individualized, clear, detailed, guiding and as immediate as possible, to allow timely intervention in learning process. However, with the short time teachers have available, it becomes complex to design, apply and evaluate formative tests (pen and paper format) in a systematic way.

This makes the use of online digital tests as a particularly interesting alternative, mainly because of the correction, grading and production of effective feedback, which can be immediate and automatic.

Online digital tests have several advantages over pen and paper tests. An online test enables the integration of multimedia elements and interactive simulations, which can facilitate a better understanding of the questions and feedback. Once the test is created, it can be automatically shared with other teachers and used by students at any time, any place and at their own pace. Students may repeat the test whenever they want, and the system can automatically shuffle the order of questions and answers on each trial. As each student finishes an attempt to solve the test, he/she immediately receives the correction, the grade and complete feedback. As all data is stored, both student and teacher can carefully analyse all the answers and the feedback for each question and in each attempt. In addition to the individual results, the teacher can analyse the overall performance of different classes and act immediately, adjusting his/her teaching strategies.

Through the use online digital tests systems, the feedback can become more immediate and effective, because it can be delivered right away and at the pace of each student, allowing him/her to understand his/hers own learning difficulties and problems and to immediately act in order to solve them (Hattie & Timperley, 2007; Stobart, 2008).

However, in the Portuguese educational context, it is not exactly known how ICTs are integrated into formative assessment activities, especially in tests. Despite the technological evolution at schools, it is considered that formative, as do summative tests, tend to be mainly applied through the traditional format - in paper.

Under these circumstances, a study was developed to identify whether teachers use tests for formative assessment tests, and, if they do not use them, the reasons and constraints they identify. In this study it was also intended to identify the possible benefits of the digital format tests for overcoming the identified constraints.

Research Aims

This work was part of a wider study focused on how online digital test systems can support effective formative assessment practices. The main objective of the study was to present a digital solution that answers to the difficulties experienced by teachers in the regular implementation of formative tests, considering the innovative technologies currently available and characteristics and functionalities that teachers value most for this purpose.

To achieve this aim, the following goals were defined: (i) identify the conceptions and the difficulties in the practices of formative assessment described by teachers, (ii) identify the specific characteristics and functionalities that teachers value most in online digital tests, (iii) identify useful features of adaptive technologies that can be used for formative assessment, and (iv) check the availability of free online platforms that match

the previous requirements identified in goals ii and iii and that are available in Portuguese and in accordance with the Portuguese educational system. If no system meets these criteria this research project intends to design a new Adaptive e-Assessment System, an adaptive online digital testing platform (v).

Methodology

Methodologically, the work was structured in three distinct phases. In a first phase, that took place between 2014 and 2015, the intention was to ascertain the concepts and practices of Portuguese primary and secondary school teachers in relation to the formative assessment tests in both paper and digital formats. Following a quantitative methodological approach, data were collected through the application of an online questionnaire. The questionnaire consisted of 14 items in the 5 points Likert scale format, ranging from 1 (totally disagree) to 5 (totally agree). Six items were about teachers' practices and conceptions on formative assessment tests and eight items about the importance given by teachers to the features of online digital tests. The discordant frequency was constituted by the congregation of the answers of levels 1 and 2 of the scale. The concordant frequency was constituted by the congregation of the answers of levels 4 and 5 of the scale. The results of level 3 of the scale were assumed as neutral opinions. Once the constraints of the paper format tests, as well as the most relevant characteristics of the digital format tests were identified, the second phase of this work was started by reviewing the literature on innovative particularities resulting from technological developments in the field of adaptive technologies for learning assessment. This process took place between 2015 and 2016. After some characteristics and functionalities of adaptive technologies had been identified, we checked and analysed the availability of free online systems that incorporate them. In a third phase, that took place between 2016 and 2017, and due to the impossibility of finding a free and appropriate platform for the purposes of this study, it was considered necessary to conceptually create and develop a new platform. This system would implement the characteristics and functionalities pointed out by teachers as more relevant and, through its adaptive particularities, should adjust to the needs of each student and teachers. The results of these three phases are described next.

Results

Phase 1

At this phase of the study a questionnaire was distributed to 793 teachers from middle and secondary Portuguese public schools. About 2% of the respondents had less than 8 years of teaching, 17% between 8 and 15 years, 45% between 16 and 25 years and 36% more than 25 years of teaching. In examining what constitutes teachers' practices regarding the use of tests in formative assessment, the results revealed that 83% of teachers apply formative assessment tests; 78% of these only use the paper format. Concerning the feedback, 34% of the teachers who apply online digital tests use tools that don't trigger it. Regarding the frequency of application of formative tests, only 18% use formative tests regularly (weekly or once every two weeks), the remaining respondents conducted them monthly or even less frequently.

Table 1

Teachers' Practices and Conceptions on Formative Assessment Tests (*n*=793).

	Agree	Disagree	Neutral
Grading and correction of tests is time consuming.	66%	12%	22%
Application of tests disrupts the teaching process and take time away that is precious to cover all the topics of the curriculum.	30%	47%	23%
Tests correction and feedback are only made orally and for all students simultaneously.	62%	20%	18%
Elaborate feedback is given to students for each question.	45%	33%	22%
The processing and analysis of tests requires too much time.	59%	18%	23%
If the correction, grading and feedback were quickly given to the students the result of applying formative tests would be more beneficial.	70%	14%	16%

As Table 1 indicates, the results reveal that teachers consider the activities of test correction and grading, as well as the provision of elaborate feedback, to be very time consuming. Surprisingly 30% of teachers consider that tests disrupt the teaching process, taking away time that is very important to cover all the topics of the curriculum. The feedback provided to students is neither immediate nor effective as it tends to be based only on the identification of the correct and the wrong answers. Most of the teachers consider that if the correction, grading and feedback were more quickly given to the students, the result of applying formative tests would be much more positive.

Table 2 presents the characteristics and functionalities teachers value most in online digital tests. These results suggest that, in teachers' opinion, the key features of online digital tests that indeed are important for formative assessment are not only the automatic presentation of feedback to the students but also the possibility of integrating multimedia elements, as it can be seen in the following table. They support the study's relevance, suggesting that online digital tests are, indeed, seen as a valid and advantageous alternative compared with the paper format, presently used by most of the teachers.

Table 2

Importance Given by Teachers to the Features of Online Digital Tests (n=793)

	Agree	Disagree	Neutral
Automatic correction and classification.	85%	4%	11%
Feedback is automatic and immediate.	89%	2%	9%
Allows effective feedback.	81%	5%	14%
Students may repeat the test how often they want.	76%	8%	16%
Allows the integration of multimedia elements.	87%	2%	11%
The integration of multimedia elements improves learning.	85%	3%	12%
Tests can be used by students at any time and any place.	82%	4%	14%
Tests and questions can be automatically shared with other teachers.	84%	3%	13%

Phase 2

Initially in this phase, and in agreement with a literature review on this topic, some characteristics of the adaptive technologies considered important to e-Assessment Systems were identified. In learning assessment, these technologies have had many benefits, especially the possibility of allowing students to use these systems autonomously and meaningfully. Among other features, the system can adapt in real time the activities and the resources provided to the students considering their needs. Based on the student's difficulties and on their answers to previous questions, the system decides what activities will be presented next. Automatically, and without human intervention, the system analyzes the student difficulties, his/her interaction with the system and the previously obtained results to suggest new learning activities adapted to his/her abilities and difficulties.

Identifying some key features of adaptive technologies, we checked and analysed the availability of free online systems that already incorporate those features. The process of searching for existing systems was based on the application of a search equation in the Google search engine, using the following keywords *adaptive learning* or *adaptive computerized testing* or *adaptive computer testing* or *adaptive e-assessment* or *adaptive computer assessment.* These were organized in a conjunctive logic with the keywords *platform, software* and *system*. For each search in Google the results of the first 15 pages were analyzed. The data extracted from the online search pointed out to a total of 25 platforms that fulfilled the defined requirements: first, to be at least partially free of cost and, second, not to be only an LMS. The 25 systems analyzed were the following: acrobatiq, Aleks, aNewSpring, area9 Learning, Cerego, CK-12, CogBooks, Concerto, Desire2Learn, DreamBox, FastTest, Fishtree, Formative, Khan Academy, Knewton, Learnetic, Learnosity, OWL, Realizeit, ScootPad, Smart Sparrow, Socrative, Surpass, TAO Community Ed., Testive.

For the analysis of these platforms, the following characteristics were considered: (i) it is an adaptive system, (ii) it has free licensing, (iii) the interface is configurable for the Portuguese language, (iv) it is simple to use, (v) it allows the integration of teacher authoring resources and questions, (vi) it provides an elaborated feedback to the student (immediate, descriptive and guiding), (vii) it allows the definition of difficulty levels for questions and tests, (viii) it provides automatic activities to students, (ix) it is possible to align it with new curriculum structures (Portuguese), (x) it is multidisciplinary, and (xi) it provides reports of difficulty and performance.

The analysis, presented in Table 3, was based on the information provided on the product websites, on white papers published by the authors of the systems and on some tests that were carried out (when allowed by the platforms).

Table 3

Analysis of the Platforms (x Means the Characteristic is Verified; N.O. Means the Characteristics is Non-Observable)

	i	ii	iii	iv	v	vi	vii	viii	ix	x	xi
acrobatiq http://acrobatiq.com	x			x	x	x				x	x
Aleks www.aleks.com	x			x	x	x		x		x	x
aNewSpring www.anewspring.com	x			x	x					x	x
area9 Learning http://area9learning.com	x			n.o.	x	x				x	x
Cerego www.cerego.com				n.o.	x					x	x
CK-12 www.ck12.org		x		x	x					x	x
CogBooks www.cogbooks.com	x			x	x	x				x	x
Concerto http://concertoplatform.com	x	x			x	x				x	x
Desire2Learn www.d2l.com				x	x					x	x
DreamBox www.dreambox.com	x			x		x				x	x
FastTest www.assess.com	x			n.o.	x	n.o.				x	x
Fishtree www.fishtree.com	x			x	x	n.o.		x		x	x
Formative http://goformative.com				x	x					x	x
Khan Academy https://pt-pt.khanacademy.org	x	x	x	x				x		x	x
Knewton www.knewton.com	x	x						x		x	x
Learnetic www.learnetic.com	x			x	x					x	x
Learnosity www.learnosity.com				x	x					x	x
OWL www.owlts.com				x	x						x
Realizeit http://realizeitlearning.com	x			n.o.	x				x	x	x
ScootPad www.scootpad.com	x			x	x			x		x	x
Smart Sparrow www.smartsparrow.com	x				x					x	x
Socrative www.socrative.com		x		x	x					x	x
Surpass www.btl.com/surpass				x	x					x	x
TAO Community Ed. www.taotesting.com	x	x		x	x	x				x	x
Testive www.testive.com				x	x					x	x

The analysis led to the following conclusions: (i) from the 25 platforms that were analyzed about two-thirds weren't adaptive; (ii) only four had a totally free license, and only two allowed implementation of assessment activities;

(iii) only one presented a native interface in Portuguese, but it was a closed system with proprietary content; (iv) most systems didn't allow the sharing of resources created by the users; (v) only seven platforms enabled elaborated feedback, but none of them enabled the definition of difficulty levels for the questions; (vi) only one allowed the alignment with new curricular structures; (vii) all systems were multidisciplinary; and (viii) none of the systems allowed the autonomous recommendation of assessment activities.

Therefore, it was considered that none of the platforms analyzed fully satisfied the functionalities and the adaptive features that are considered to be essential for an Adaptive e-Assessment System. Due to the impossibility of finding a free platform that could be adapted to the Portuguese educational system, and that suited the purposes of this work, the study evolved to its third phase.

Phase 3
As no system had been found that meets the requirements defined in the previous phase, the conceptual design of a new Adaptive e-Assessment System started to be developed. For this process the first step was to design the framework that supported the system. The basis of the framework is presented in Figure 1: (i) Learner, (ii) Teacher, (iii) Adaptive Engine, (iv) Test, (v) Instructional Resource, (vi) Question and (vii) Curriculum. The Learner and the Teacher dimensions represent the users who will interact with the system. These as well as the other dimensions are described below.
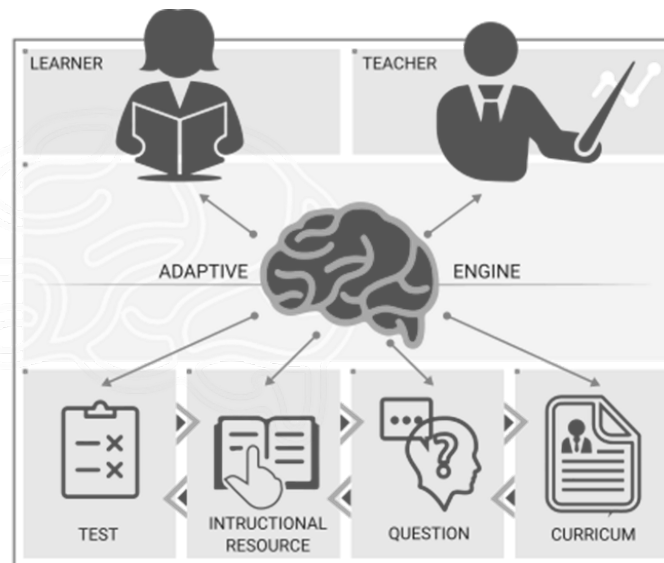

Figure 1. Framework for Adaptive e-Assessment System.

Dimension learner. This dimension represents the model composed by detailed information of the academic data and profile of each student, namely: learning style and preferences; level of general difficulty; level of difficulty per curricular topic; activities carried out and results; progress and trends; proficiency regarding each topic of the different subjects of the curriculum; interactions and telemetry; temporary records of system use; among others. This dimension allows the Adaptive Engine (AE) to make more assertive activities recommendations according to the knowledge got from the student overall activities.

Dimension teacher. The teacher plays a crucial role in system, for his or her knowledge on pedagogy, curriculum, evaluation processes and adequacy of learning approach to the characteristics of his/her students. This dimension represents the information that the system should record about the teacher-type user, particularly with respect to teachers' profiles, subjects, classes and students they teach, activities developed and assigned to each of their students, and all the interactions with the system. The greater the interaction with the system, the better the AE can help teachers in the design and management of their activities.

Dimension question. Questions are the key element of any adaptive assessment system. It is through the answers to specific questions that the system analyzes and decides what resources to recommend in order to help students to overcome difficulties and promote their learning. For instance, in a typical adaptive assessment system, if the student misses a question then an easier question should be presented next. This dimension represents the information about each question available in the system. This information goes through its description, formulation, typology (multiple choice, true/false, etc.), response options, its corresponding correction, feedback, difficulty level, author, among others. The model also defines the articulation between the *question* dimension and the remaining ones. For instance, with the dimension *curriculum* it allows categorizing each question with the topics of the curriculum to which it refers; with the dimension *test* it permits registering in which test each question is integrated; and with the dimension *instructional resource* it allows defining what resources are presented to the student together with a given question.

Dimension instructional resource. This represents each learning object available in the system, which is composed by information that relates it to a specific subject matter. This model represents the set of information that integrates each learning object, which can be made up of different formats (text, image, audio, video, interactive simulations), and is related to a certain topic of the curriculum. For example, a text, a video, a simulation that describes the respiratory system. It can be used as part of a test, articulating with the dimension *test,* or it can precede a certain question, or even be integrated in feedback.

Dimension test. A test consists of a set of questions, with a well-defined organization and structure with specific objectives in knowledge assessment and promotion of learning. It can also be supplemented with instructional resources that help to contextualize questions. This structure represents the information related to each of the tests available in the system, such as the questions and instructional resources that form it, i.e., the sequence, the level of difficulty, the total amount of time required, who was it appointed to and who solved it, as many other information.

Dimension curriculum. This represents the content of each subject, which is modelled as a hierarchical and sequential structure of knowledge items. Each item represents atomic knowledge about a particular topic. This dimension represents the detailed information of the curriculum of each subject, such as

the description of each topic, its sequence and hierarchy. To have mastery over a certain subject, the student must dominate all its elementary parcels - the items. It is with these items that the questions and instructional resources are articulated. Thus, the AE can understand what questions and resources are available for each curricular topic. The curriculum dimension represents this information in detail and also structures the interdependency of the topics.

Dimension adaptive engine. This dimension represents the model responsible for the interconnection of all components of the system. It is the central element of connection and communication among all other dimensions and has the function of learning with and about the users, the questions, the instructional resources and the curriculum. It represents the characteristics, functionalities and adaptive algorithmic mechanisms that the system must implement to accomplish its functions efficiently and effectively. It must discover and adapt the activities to be carried according to the profile, behaviour and difficulties of each student, assuring him or her an individualized learning process. It manages all processes, interactions and presents the right information at the right time and place. It ensures that the registration, updating and analysis of all the information flow in the system in real time, at the same time that keeps the entire system functioning.

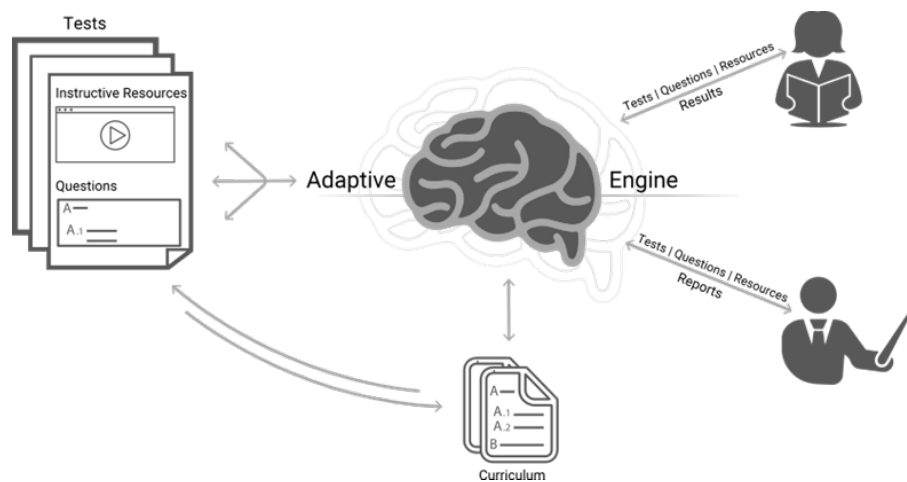Figure 2 presents a diagram of the articulation between these dimensions.

Figure 2. Articulation between the framework dimensions.

Conclusion

Therefore, the next stage of research, currently under development, consists of the conceptual design of a new Adaptive e-Assessment System that responds to the needs of the teachers and contributes to the students' learning. The system, which we named *TRAiL-Transforming Assessment into Learning*, will be based on the framework previously presented.

It is considered that the instructional model "one-size-fits-all" is not suitable to the diversity of learning profiles, needs and motivations of todays' students.

Thereby, the system allows all assessment activities to be adjusted to the characteristics, difficulties and capacities of each student.

In general, it is assumed that TRAiL will include the following particularities:

- Available and suitable for all subjects and years of schooling
- Availability to be used at any time, place and device
- Allows the student to perform the same activity as often as he/she wishes
- Allows the teacher to characterize the profile, capacities and difficulties of each student in the system
- Allows the teacher to create resources, questions and tests, that will be automatically shared with other teachers
- Allows the possibility of student to perform activities even without the intervention of the teacher, either on his/her own initiative or by suggestion of the system
- Guarantees access by the teacher to the data of his students, and the access of students to their own data, enabling the monitoring of the results obtained by them in all activities
- Has a diversity of types of questions with the possibility of integrating multimedia elements in their formulation and response options
- Allows feedback in multiple media formats (e.g., audio or video)
- Includes elaborated feedback mechanisms (immediate, descriptive and guiding) for each response option, and provides the student with content which he wouldn't probably have contact with on his/her own studying activities
- Has the possibility of creating and suggesting autonomous tests to the students, considering the resources and questions available on the system and according to different levels of difficulty and performance
- Analyzes in real time the difficulties and performances of the students, as well as the results and performances available for each of the questions and tests
- Allows a systematic analysis of the levels of difficulty and results in each topic of the curriculum
- Enables the creation of automatic tests either by teacher or student in a reduced number of steps, through the reuse of questions and resources available in the system
- Has the flexibility to support the development of other types of assessment (diagnostic and summative) and to be used in classroom activities (e.g., classroom questioning and discussion)
- Has the ability to learn from their own analysis, decisions and actions

Thus, we consider that the future development of TRAiL for the Portuguese educational system will contribute to a regular use of formative assessment tests that develop students' selfregulation and academic success.

Acknowledgements

References

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan,* 80(2) 139-44.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning.* Thousand Oaks, CA: Corwin Press.

Gilbert, L., Whitelock, D., & Gale, V. (2011). *Synthesis report on assessment and feedback with technology enhancement*. Retrieved from https://core.ac.uk/download/pdf/ 1512614.pdf

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Irving, K. (2007). Formative assessment improves student learning. *NSTA Reports, 18*(7), 6-8.

Marzano, R. (2007). Designing a comprehensive approach to classroom assessment. In. D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 102–125). Bloomington, IN: Solution Tree.

Roediger, H., Agarwal, P., McDaniel, M., & McDermott, K. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied, 17*, 382–395.

Roediger, H., & Butler, A. (2013). Retrieval practice (testing) effect. In H. Pashler (Ed.)*, Encyclopedia of the mind (pp. 661-663)*. Los Angeles, CA: Sage Publishing Co.

Stobart, G. (2008). *Testing times*. Abingdon, United Kingdom: Routledge.

Author Details
Nuno Dorotea
nmdorotea@ie.ulisboa.pt

Neuza Pedro
nspedro@ie.ulisboa.pt