

International testing platforms: When similar outcomes meet different testing needs

Jesús García Laborda

Universidad Politécnica de Valencia, Grupo GILFE, Carretera Nazaret - Oliva, s/n, 46730 Grao de Gandía (Valencia), España

Computers have become a solution for many traditional problems from the kitchen to the school. It is not, therefore, uncommon to use them to solve existing problems. It is also quite frequent that different educational needs can lead to similar solutions. For instance, the Australian immigration services use a number of exams to allow the access of qualified professionals into the country. Likewise, the TOEFL exam is commonly used as a requirement for foreign students to access to American universities. Similarly many European universities that receive a large number of international students every year have developed their own language test to admit or place foreign students.

The Universidad Politécnica de Valencia is probably among the top three universities in Spain by the incoming number of international students. Only in the last three years the Gandía College (UPV) has tripled the total number of these students. Thus, creating and developing especially tailored courses for them is quite a complex issue. One of the reasons is the content teachers' limited knowledge of foreign languages but also, probably the most important, the difficulty of having reliable information of the international students proficiency level upon arrival. Many universities use their own language services to assess the national and international students but these tests are usually free of charge for the students and are costly in human and economic resources, and consequently many universities tend to use very limited resources to diagnose the incoming students. In addressing this matter, GILFE, a research group in languages and technology at the Universidad Politécnica de Valencia, has been studying this problem for three years and two years ago started a research project to develop a computer based test to do this job (García Laborda, 2004; García Laborda & Bejarano, 2005). This computer tool was thought to be Internet based and would be available for low stake testing and is currently in trial.

Palabras clave testing platforms; exams; TOEFL; interface design

1. General Introduction

Computers have become a solution for many traditional problems from the kitchen to the school. It is not, therefore, uncommon to use them to solve existing problems. It is also quite frequent that different educational needs can lead to similar solutions. For instance, the Australian immigration services use a number of exams to allow the access of qualified professionals into the country. Likewise, the TOEFL exam is commonly used as a requirement for foreign students to access to American universities. Similarly many European universities that receive a large number of international students every year have developed their own language test to admit or place foreign students.

The Universidad Politécnica de Valencia is probably among the top three universities in Spain by the incoming number of international students. Only in the last three years the Gandía College (UPV) has tripled the total number of these students. Thus, creating and developing especially tailored courses for them is quite a complex issue. One of the reasons is the content teachers' limited knowledge of foreign languages but also, probably the most important, the difficulty of having reliable information of the international students proficiency level upon arrival. Many universities use their own language services to assess the national and international students but these tests are usually free of charge for the students and are costly in human and economic resources, and consequently many universities tend to use very

limited resources to diagnose the incoming students. In addressing this matter, GILFE, a research group in languages and technology at the Universidad Politécnica de Valencia, has been studying this problem for three years and two years ago started a research project to develop a computer based test to do this job[1][2]). This computer tool was thought to be Internet based and would be available for low stake testing and is currently in trial.

However, although the needs were somehow different, as soon as the first version of the new iBT TOEFL appeared (see <www.ets.org/toefl> for further information) the similarities were self evident. This paper will try to present the tool and also show the differences between the two platforms. It is not in the scope of this paper to address the differences of content and testing procedures between them but to point out the concept differences between both testing platforms. There is little question that further research is necessary in relation to the development of new language platforms and, in this sense, this paper is only a tentative approach to such an attractive issue.

2. Integrating skills

Given the basic and initial needs to develop PLEVALEX (Platform of Evaluation of Foreign Languages), it was self evident the need to develop a tool that could integrate all the skills that students would need in their social and academic lives. It seems unnecessary to mention that all four skills should be integrated in the computer platform. These skills should be integrated in realistic environments. It should also be relevant to the students' realities and needed to include certain types of oral and written exercises that could replace the traditional pen-and-paper and pair-interview tasks that were so common until the beginning of the 2004-2005 academic year. The incorporation of academic mini-lectures in the exam made also think the developers the possibility of incorporating note taking as a need. Up to here, the coincidences with the new iBT TOEFL changes are also clear. The inclusion of these tasks responds to the fact that the traditional tests and oral exercises had a very limited relationship with the university routines that include attending classes, writing well planned papers, take exams of out-of-classroom readings but also in-class lectures. The exams distributed by PLEVALEX seem to ensure a fair but also useful diagnosis.

What is the internal organization?

As opposed to the TOEFL, the PLEVALEX platform allows the student a great flexibility. This paper will not describe each section in depth that can be found in other studies (García Laborda, 2006) but it is also important to mention that the examinee can choose in which order to proceed when interacting with the index: 1) grammar / reading (HIELE), 2) writing (ESCRITOR), and 3) speaking (HIEO). HIELE uses multiple choice questions mostly, ESCRITOR includes as many writing essays as required (usually 1 or 2), and HIEO has short social or academic questions and longer descriptions. It is mostly in this last section where students will be freer to use notes taking at the time of playing the videos in the speaking section. TOEFL has four sections but will be done in order: reading, listening, speaking and writing¹. PLEVALEX designers, researchers and administrators strongly believe as those in TOEFL that obtaining a threshold grade in the tests distributed with these two platforms can be considered reliable as the tests will match most academic demands required in higher education. It is self evident in both platforms that communication has overtaken grammatical accuracy and language use, and that criticism of the validity of language testing as a predictor of language proficiency can be clearly override. This position is strongly supported because the idea of PLEVALEX's designers is to achieve a flexible and adaptative test in a near future and being able to equate each student against the Common European Framework (CEF) instead of having the currently developed norm based exam at B2/3 proficiency level. From the overall diagnosis based on this proficiency level, it is expected to obtain valuable information to be used

¹ Specimens of the different types of tasks that are mentioned in this paper can be found in the website mentioned above likewise in [3] the students will also find a complete selection of interfaces used in the exam.

by the different institutions across Europe and also provide feedback for international student mobility preparation (washback).

Exam sections

The new testing platform can deliver a competitive test that can measure the international students arriving to that university. To make a short description, the tool is composed by three modules: written composition, multiple choice and a speaking skills sections.

Section →	Multiple Choice	Written Composition	Speaking section
Tasks →	Choose one (correct) answer	Short answers Prompted or figure aided compositions	General pair directed questions Lecture and question answer Diagram / picture description

Fig.1: Exam sections

As mentioned before, the platform integrates all four skills and follows the same general structure

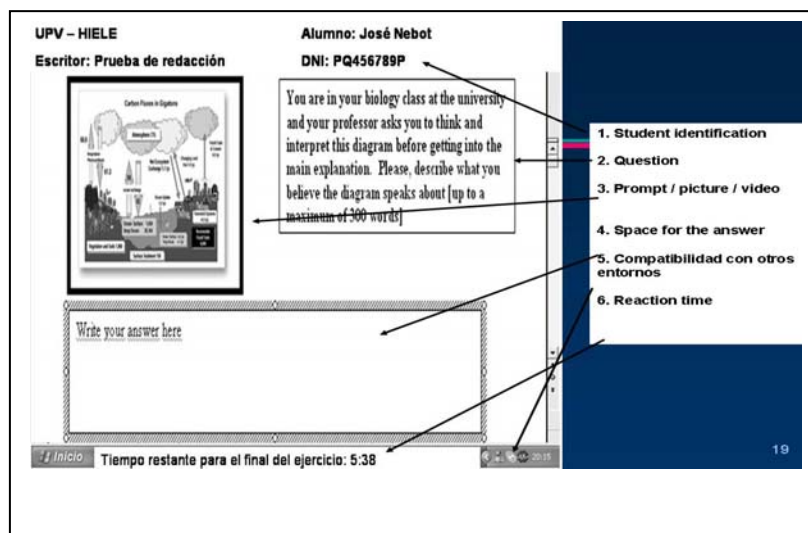


Fig.2: Interface design for exams. Please note the different parts of the composition section

This structure is very simple and tries to avoid any possible interference due to any stimuli not directly related to the exam and the minimum requirements to represent the prompts, question of space for response. The main difference in the design stage is that while PLEVALEX uses limited elements specifically developed for this platform, all the elements included in TOEFL were specifically design for their test. Of course, this has to be with the limited budget that the Spanish platform had totalling 24,000 Euro.

This schemata was used for the writing sections mostly by both platforms and, in relation to PLEVALEX, it was also used for the speaking section without the space for writing and being substituted by an ON/OFF speaking icon.

In relation to the “language use” / grammar, both platforms used single question interfaces.

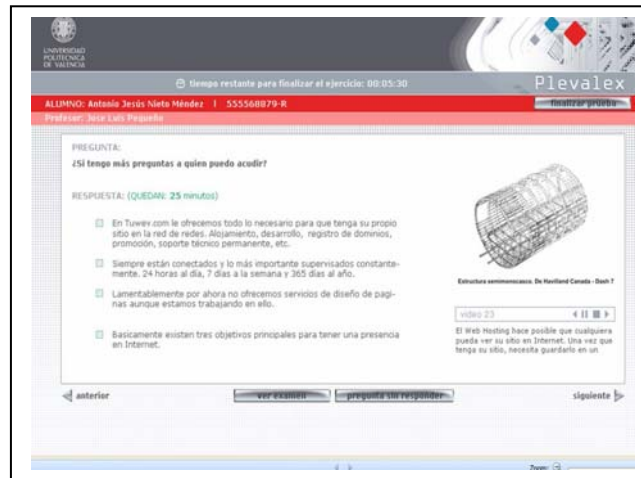


Fig. 3: Multiple choice or reading section

Other than that and as it has been made clear by the interfaces shown in this short paper are clear so, how does PLEVALEX outcome TOEFL? Is it worth to develop platforms like PLEVALEX?

Possibilities for PLEVALEX and further developments

It is self evident that the limitation of author tools like TOEFL makes them extremely expensive for low-stakes exams. Universities could produce tools like this to evaluate large quantities of students when the results can have a limited scope for the examinees. For example, the Universidad Politécnica de Valencia will be using and testing this tool in the Foreign Language Institute (that means, as opposed to TOEFL, at least in 6 languages) in the first semester of 2006. Usability tests with Spanish and international students totalling 35 reports and questionnaires have proved that the platform is easy to use and comprehend. Further studies should consider how to use this platform for high stakes exams like the National Entrance Examination, defining a theory of language test interface and platform design, students' differences due to the use of keyboard (versus pen and paper), individual strategies, and many more.

Plevalex and TOEFL face to face

When preparing the test and platform, the researchers intended to compare this exam mainly to the new version of the TOEFL to which the exams distributed through HIEO (Herramienta informática de Evaluación Oral) and HIELE (Plataforma de Evaluación de Lenguas Extranjeras) resemble. The following diagrams in the poster summarise the differences between both platforms.

Conclusion

The PLEVALEX platform has an enormous potential as it can integrate all four skills plus grammar in a controlled and easy-to-distribute test. It also helps to show semi-communicative skills and situations. Its format also benefits the using institutions as it allows a better use of the testing budget by using it for raters the more human tasks like grading speaking and writing. Although this complex project still remains unfinished the short and long run perspectives are very positive. The prospective studies in examinee satisfaction and comfort will certainly bring new improvements to this software.

References

- [1] J. García Laborda "HIEO: Investigación y desarrollo de una Herramienta Informática de evaluación oral multilingüe" *Didáctica. Lengua y literatura*, 16: 77 – 88. (2004)

