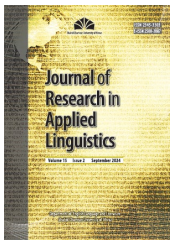
Journal of Research in Applied Linguistics
ISSN: 2345-3303 – E-ISSN: 2588-3887 – https://rals.scu.ac.ir

Please cite this paper as follows:

Arrosagaray, M., & Urreizti, G. (2024). Towards digitalised language assessment in Spanish official language schools (EOI): Current opportunities and challenges. *Journal of Research in Applied Linguistics*, *15*(2), 89-102. https://doi.org/10.22055/rals.2024.45500.3207



Research Paper

Towards Digitalised Language Assessment in Spanish Official Language Schools (EOI): Current Opportunities and Challenges

Marcelino Arrosagaray[1] & Garbiñe Urreizti[2]

[1]Corresponding author, Departamento de Ciencias Humanas y de la Educación, Facultad de Ciencias Humanas, Sociales y la Educación, Universidad Pública de Navarra, Pamplona, Spain; *marcelino.arrosagaray@unavarra.es*

[2]Departamento de Ciencias Humanas y de la Educación, Facultad de Ciencias Humanas, Sociales y la Educación, Universidad Pública de Navarra, Pamplona, Spain; *garbine.urreizti@unavarra.es*

Received: 06/01/2024 Accepted: 24/09/2024

Abstract

Online and hybrid courses that integrate technology and complement in-person teaching are increasing in Official Language Schools (EOI, Escuelas Oficiales de Idiomas) in Spain. Digitalised online assessment, as a process encompassing taking the tests, either in person in the traditional classroom or remotely and subsequently their correction and marking, renders the logical outcome of this technology-mediated learning. In order to analyse the digitalisation process and the adaptation of our traditional print-based tests to a new computerised format, at the Distance Official Language School of Navarra (EOIDNA), one of the three EOI in the province of Navarra, we have carried out an experimental study with students who took their language tests (N=144) either under direct in-person instructor supervision (N=127) or through proctored online software (N=17). In this paper, we describe the contributions generated as perceived by students and test instructors. Results have shown that a vast majority of students reported a positive or neutral effect of taking the test on their computers and stated that the experience was satisfactory. Additionally, the transformed tests have resulted in quickly accessible and flexible administration and efficient management and storage of student test records for instructors.

Keywords: Digitalisation; Language assessment; Official Language Schools (EOI) tests; Student attitudes.

1. Introduction

These days learning a language is neither exclusively related to a moment, compulsory education, nor a place, school or university. The intensive presence of languages in the Spanish curriculum of Primary, Secondary and Tertiary education and the rapid expansion of technology, which has allowed easy access to sources of information and entertainment in the target language itself, have contributed to fueling social interest in language learning in new emerging formats.

In order to meet this interest, fully online or hybrid courses that integrate technology and complement traditional classroom practice have increased in Spanish EOI. These courses provide students with flexible and personalised learning opportunities that can be adapted to their needs and demand the students' active engagement. In this vein, digitalised language assessment as a process encompassing taking the tests, either in person in the traditional classroom or remotely, and subsequently, their correction and marking, is the logical outcome of this technology-mediated learning.

Currently, educational administrations are incorporating computerised tests in their diagnostic or system evaluations, such as PISA (Program for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study) or PIRLS (Progress in International Reading Literacy Study), which are mostly administered by computer. As previous research has defended, technology seems to facilitate assessment implementation and maximize the benefits for all involved stakeholders (Timmis, et al., 2015; Laborda et al., 2015; Okada, 2019). However, its use is not yet widespread, although it can be affirmed that access to the Internet and digital devices is contributing to the increase

in computer tests (Suárez-Álvarez et al., 2022). In tertiary or post-Secondary education there seems to be also limited implementation of online invigilated examinations (Butler-Henderson & Crawford, 2020).

Due to COVID-19 pandemic, new online learning and teaching scenarios were generated, among which online assessment was included. This experimentation, which forced us to innovate urgently, highlighted the shortcomings of the educational system and made it possible to open new ways of teaching and assessment. For the time being, it is our belief that these new scenarios need to be refined ad consolidated with conviction if they prove effective and desirable.

Educational transformation usually arises because of normative regulation. On many occasions, this normative regulation is not fundamentally based on experimentation, technical viability or the demonstrated educational effectiveness of the new measures, but rather on the assumption of sociological trends or the obligatory application of higher-level regional or national regulations. These innovations or new ways of operating, mostly imposed on a top-down basis, do not necessarily imply a pedagogical improvement in terms of learning or teaching.

In this line, online assessment can be considered non-existent in Spanish EOI up to date, to the best of our knowledge. Thus, our study intends to put forward the acquisition of educational criteria based on the required experimentation and justification to promote a profound educational change in our environment. This transformation must guarantee the quality of the assessment itself and the feasibility and proper management of the possible technical complexities derived from its implementation. The quality of online assessment might then rely on the suitability of this analysis, as there might be several factors involved in test improvement policy (Darabi & Ahmadi, 2019).

The consolidation of digital teaching and learning processes is one of the objectives of our educational institution. EOIDNA currently has a considerable number of students who follow distance or hybrid courses in English, Basque, French and German (listed in order of greatest participation): Secondary, Baccalaureate and Vocational Training students who are enrolled in studies in Navarra or abroad temporarily (N=5766), university students (N=83) and adult students of various origins through the *That's English* Program for adults (N=1068).

Thus, as a publicly funded Official School of Languages, we carried out a pilot study to analyse the institutional action needed for the implementation of digital assessment, the technological equipment necessary to carry out the examinations digitally, the possible influence of digital tests on student academic performance, and finally, students' perceptions after the experience.

2. Theoretical Framework

2.1. Adults' Language Learning and the Spanish EOI

To understand the nature and scope of Spanish EOI, we will refer to the European Union language learning policy, the amount of the Spanish population under the tuition of Official Language Schools and the type of courses offered in them. One of the objectives assumed by the European countries in their educational policy ET2020 (Education and Training 2020) was the participation of at least 15% of the adult population in lifelong training programs. In a broad context, figure 1 below shows the evolution of participation in a comparative way between Spain (green line) and the EU (with 27 members, as of 2020 due to Brexit) and with 28 members until then (light and dark brown lines, respectively).

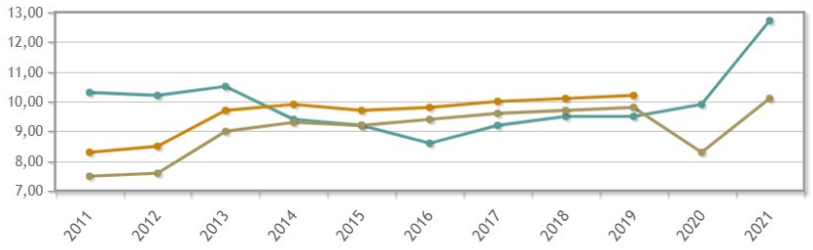


Figure 1. *Adult Population in Lifelong Training Programs*
(Source: European Commission (2023). Key data on teaching languages at school in Europe)

We can affirm that adult citizens' participation in formal or informal training in Spain is comparatively greater than that of other European countries, although it does not reach the desired objective. Within Spain, the statistics for the year 2021 show the following comparative results by autonomous community:

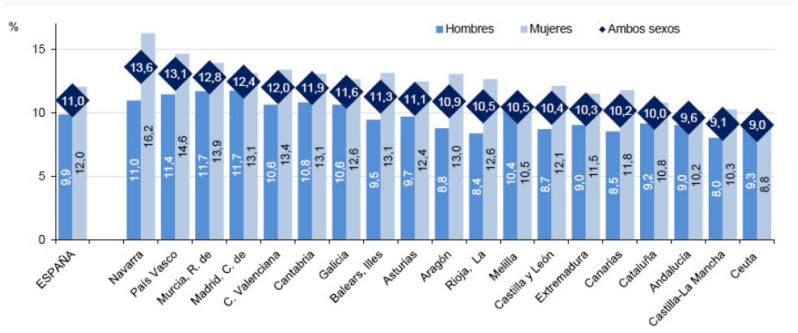


Figure 2. *Adult Population in Lifelong Training Programs in Spain by Autonomous Community* (Source: Consejo Escolar del Estado (2022). Informe sobre el estado del sistema educativo).

As stated in the Annual Report on the State of the Educational System, *Informe sobre el estado del sistema educativo* (Consejo Escolar del Estado, 2022), the COVID-19 pandemic slowed down the already slow progress of adult learning throughout the European Union, although it did not alter the low average participation rates or the unequal image in the Member States. Nonetheless, the pandemic has generated momentum in adult learning as a policy objective. As the OECD report Policy Responses to COVID-19 (2020) highlights, large-scale publicly subsidized distance learning provisions or support measures concerning adult education and training are scarce. About 28% of adults claimed they did not participate in training because they lacked time due to work commitments and another 15% reported a lack of time due to family responsibilities. An additional 16% mentioned a lack of financial resources and 12% stated that training took place at an inconvenient time and place (OECD, 2020: Policy Responses to COVID-19).

In the 2020-2021 academic year, as shown in the aforementioned national report (Consejo Escolar del Estado, 2022), Spain had 392 centres for language teaching between Official Language Schools (our EOI) and Affiliated Classrooms (*aulas adscritas*). The total number of these schools, run by the state and distributed throughout the different Spanish Autonomous Communities and cities, teach language programs for adults. In turn, the students enrolled in the EOI in Spain were 380,278 (which represents approximately 1.01% of the total population of the country), of whom 346,286 followed their classes in person (91%) and 33,992 did so entirely online (9%), according to the same source (Consejo Escolar del Estado, 2022).

As can be seen, this public educational offer of the Official Language Schools (EOI) is mainly materialized in face-to-face courses. These cover all language levels ranging from A1 to C2, considering the structuring of language proficiency levels of the Common European Framework of Reference for Languages (CEFRL). They consist of 130 hours of attendance in the classroom for students aged 16 or above as a general requirement. In addition, schools are increasingly offering the possibility of a blended or hybrid learning modality, replacing part of this direct teaching with remote work through platforms or virtual learning environments. The students enrolled in these two types of courses are called *official students*. Alternatively, candidates may opt directly for the certification tests through basic registration without any kind of academic guidance. This is the case of the so-called *free students*.

In turn, the distance education offered in the Spanish EOI is mainly concentrated in the English teaching *That's English* program, run by the Ministry of Education and Vocational Training. This program offers the possibility of attending face-to-face tutorials once a week, which is complemented by a platform or virtual learning environment (e-learning) that has interactive and self-correcting activities with multiple teaching aids so that students can study English autonomously and enjoyably. So far, this program covers most language levels, from the Basic A1 level to the Advanced C1 level. In addition to the course English activities, this program offers access to extra content, a podcast, and a complete online English dictionary.

Lately, in order to further diversify their educational distance learning offer, the *That's English program* has led to the production of a series of fully online preparatory courses of certification exams. These courses, structured by levels

from A1 to C1, consist of a section on strategies that train test candidates to deal more effectively with the examination tasks, a practice section and a package of five complete exams per level intended to reinforce the preparation of candidates.

As of the 2021-22 academic year, some intensive preparation courses for certification exams in the hybrid modality or blended learning, called Exam Prep, have also been developed for levels B2 and C1. These courses, offered experimentally in Navarra through our school (EOIDNA), have a reduced duration of four months (the equivalent of a total of 120 tuition hours) and consist of material for familiarization and practice in the several test parts. They also offer the possibility of holding several videoconferences with a tutor to practice oral expression in very small groups and to carry out written expression exercises that are corrected and commented on by the tutors. Enrolment data for the first two editions of the course prove its acceptance and student interest in this new format.

2.2. ICT Use in EOI Language Teaching

The incorporation of ICT and commonplace online technology into the language classroom has lately become mainstream (Stickler & Shi, 2016). It cannot be denied that ICT tools have become essential and the standardization of their use is ever more present in language learning and teaching settings, especially in distance education.

In this push for the digitalisation of education, there are three major educational gaps that educational institutions aspire to close: equitable access to technology, the quality of use of these tools, and training to develop and use them. In turn, most research on distance language courses seems to focus on technological aspects in language teaching, the effectiveness of technology in the development of linguistic competence and autonomous learning (Nurieva & Garaeva, 2020).

According to the aforementioned Spanish official report (Consejo Escolar del Estado, 2022), in 2021 83.7% of Spanish households had some type of computer or digital device (desktop, laptop or tablet) and 95.9% of Spanish households had internet access, compared to 91.4 % in the year 2019. Almost all had broadband internet access (optical fibre or cable network, ADSL, mobile telephony, etc.). The main type of broadband connection was through fixed modalities (optical fibre, ADSL or others), present in 82.9% of households, while 13% of households access was made using only mobile connections through 3G, 4G or 5G devices.

As for schools, the same report defines three different types of computers: (1) those intended for direct teaching of students and the tasks of the teachers; (2) the ones preferably intended for direct teaching of students, used mainly in the teaching of classes or for students to work practically, mainly located in computer rooms or ordinary classrooms; and (3) those intended for teachers' tasks, used by teachers to prepare classes or follow-up of the students and usually located in the rooms or offices of the teaching staff. Thus, there are 2.5 students per computer in type 1 above, 2.9 students per computer in type 2 and 1.6 teachers per computer in type 3. The internet connection of schools in Spain is practically 100% from the 2005-2006 academic year, and in the 2020-2021 academic year, 97.4% of regular classrooms had a connection.

Regarding the quality of use of ICT in our specific context, previous research (Arrosagaray et al., 2019) points to general trends also based on the learning mode of the EOI, face-to-face, blended or distance learning. Teachers are mostly present through presentations and explanations in the first scenario as compared to the other two learning modes. Also, there is little interaction among students in the blended and distance courses, compared to face-to-face instruction, with very little room to hold discussions with a view to improving their learning. So, digitized assessment practices seem to be scarce altogether too regardless of the learning mode. This research also points out that most students feel somewhat confident in carrying out ICT-related activities, with distance learning students showing the greatest degree of self-confidence.

The outbreak of the COVID-19 pandemic generated the sudden need to attend to students without being present in the classroom in the EOI, as in other educational contexts. In this adaptive effort, the evaluation of the written tests at the end of the academic year did not suffer any alteration. It was carried out in person, applying the current health and safety protocols. On the other hand, the oral tests could be carried out electronically. In our case, EOIDNA students were allowed to sit the oral test online via the Zoom application or postpone the in-person oral examination for the coming year at no additional cost. Finally, we evaluated more than 800 students online, with the consequent organizational challenge. Experience showed that online assessment could maintain sufficient quality standards. It also ensured the

implementation of existing test protocols like paired speaking, thus preserving the interactional potential of this skill, also shown in previous research (Soodmand Afshar, 2020).

In this sense, after emergency digital experiences, when surveyed comparatively sometime later, university foreign language programs show an improved level in the digitalisation of teaching resources and better digital competence levels. These programs also offer more refined professional, communicative, technological and organised peer support to face the challenges when adaptating to a digital and hybrid format (Makhachashvili & Semenist, 2022).

This seemingly consolidated educational context, in terms of logistic endowment and resistance test against unforeseeable events, brings about the emerging need for educational institutions to provide a more dynamic and productive curriculum that reflects the changing nature of society and students' literacy needs and abilities while guaranteeing educational quality and academic viability. And this should also involve assessment as part of the learning process in our view.

2.3. Digitalised Language Assessment

Language assessment constitutes a specific dimension in educational testing and entails diverse implications for test developers, instructors, and test-takers (Invar-Lourie, 2013). This field seems to have evolved together with advances in language teaching, though at a slower pace (Razavipour, 2014). Thus, however paradigmatic the use of technology might become in language teaching, it does require careful analysis of the extent to which technology is incorporated into language testing and its effects on different stakeholders.

As it has been pointed out in this paper, educational administrations are incorporating tests administered by computer into their diagnostic or system evaluations. In the field of languages, Cambridge and TOEFL also offer candidates the possibility to take computer-based tests. However, this trend is not widespread and the pace of the digitalisation of assessment seems to be slow. There is a need for organizational culture and readiness and logistical difficulties should not be underestimated either for this purpose.

To the best of our knowledge, to date, examples of digital assessment mainly include multiple-choice tests, standardized tests, digital portfolios, and proctored exams. Previous research (Timmis et al., 2015; Okada et al., 2019) points to the integrity of digital assessment processes and several associated advantages: time flexibility to apply or sit the tests, instant and more objective marking, more efficient management of assignment submissions and archiving of student attainment records, among others. In addition, the results of examination across different administration methods have suffered no variation in score (Gold & Mozes-Carmel, 2009; Butler-Henderson & Crawford, 2020). This fact itself may indeed account for their validity, whichever initial prejudice may exist for or against digital tests.

The future of assessment should address some challenges and make digital assessment authentic, accessible, appropriately automated, continuous and secure (Pauli & Ferrell, 2020). In the process of transforming EOI traditional assessment methods into digitalised ones, some further light should be shed to better understand the specific challenges involved, as is our goal in this paper.

Traditional print based EOI examinations are public standardized tests that follow national requirements as dictated by the Royal Decree 1/2019, which establishes the common basic principles of evaluation to be applied to the official certification tests of B1, B2, C1 and C2 levels. In their edition process, our EOI B1, B2, C1 and C2 level tests follow a formally established protocol: each level test is initially created from scratch by a commissioned teacher, piloted with students in a different autonomous community, their results being analyzed with *TiaPlus* statistical software, and then validated with a final revision by the whole teaching staff of each language department before they are finally edited by the commission and administered by each school. On the contrary, the assessment of A1 and A2 levels depends on regional criteria issued by the Education Department of the Regional Government of Navarra and does not follow the previous procedure but is designed and administered by each school. In all cases, the test serves as an end-of-course summative examination to validate students' knowledge and language competence in five different areas or skills in B and C levels -reading comprehension, listening comprehension, writing, speaking and (linguistic) mediation, this fifth skill not being included in A level. These tests are held in traditional face-to-face invigilated rooms. Figure 3 illustrates the different sections of each level test, their duration and passing score requirements:

Level	Written test				Oral test		Passing scores
	Reading	Listening	Writing	Duration	Speaking	Duration	
B1	4 texts 20 marks 45 minutes	4 tracks 20 marks 40 minutes	Written Mediation 100-120 words 10 marks / 30 minutes Writing 2 tasks / 60 minutes Task 1 (80-100 words) / 10 marks Task 2 (130-150 words) / 10 marks	2h 55min	Spoken mediation 10 marks Preparation: 2 minutes Speaking: 3 minutes Speaking 20 marks Monologue: 1+3/4 minutes Interaction: 1+4/5 minutes	26 minutes (2 candidates)	50% in each area & overall 65%
B2	4 texts 20 marks 60 minutes	4 tracks 20 marks 45 minutes	Written Mediation 120-150 words 10 marks / 30 minutes Writing 2 tasks / 75 minutes Task 1 (100-120 words) / 10 marks Task 2 (180-200 words) / 10 marks	3h 30min	Spoken mediation 10 marks Preparation: 4 minutes Speaking: 3 minutes Speaking 20 marks Monologue: 1+3/4 minutes Interaction: 1+4/5 minutes	30 minutes (2 candidates)	50% in each area & overall 65%
C1	4 texts 20 marks 75 minutes	4 tracks 20 marks 50 minutes	Written Mediation 140-170 words 10 marks / 30 minutes Writing 2 tasks / 105 minutes Task 1 (120-150 words) / 10 marks Task 2 (250-280 words) / 10 marks	4h 20min	Spoken mediation 10 marks Preparation: 10 minutes Speaking: 4 minutes Speaking 20 marks Monologue: 1+3/4 minutes Interaction: 2+5/6 minutes)	36 minutes (2 candidates)	50% in each area & overall 65%

Figure 3. *EOI Test Description* (Source: EOIDNA records attending official specifications)

In our view, the most important issues that the digitalisation of these EOI examinations entails include:

- Firstly, ensuring the validity and reliability of the tests, since we also believe that psychometric rigour must accompany technologically innovative tests (Hernández et al., 2021). This mainly requires having adequate software to design the test and comply with its administration regulations.
- Secondly, digital tests should guarantee students' access, satisfaction and fidelity. Digital tests should be accessible to everyone to the greatest possible extent, including those who have a long-term disability or a short-term injury. These enabling aspects should contribute to developing a positive attitude towards digital assessment and a will to try the experience in the future.
- And, thirdly, the teachers' capacity to adapt to a new digital environment. This requires different supervision and management skills to face unwanted technical problems caused by the technology used or possible dishonest behaviour from candidates' cheating attempts.

These will be the factors deeply analyzed throughout the pilot intervention carried out by our institution in order to ensure quality and maintain public confidence in qualification standards. Our study intends to address the following issues:

1. Which institutional action does the implementation of EOI digital assessment require?
2. Which technological equipment is necessary to apply EOI examinations digitally?
3. Do EOI digital tests have any influence on the student's academic performance?
4. What are students' perceptions after the experience?

3. Methodology

3.1. Context

EOI certification tests in Navarra (B1, B2, C1 and C2 level tests) are unified and simultaneously administered. The same tests are administered on the same day and time in the three official schools of the autonomous community.

As a public institution, it became necessary for our school to get the formal approval of educational authorities in the province to modify the assessment modality and their technical support to ensure its viability. The initiative was first put on the table by the EOIDNA direction board at the beginning of the academic year and the first contacts were established with the educational authorities about nine months in advance. It was assumed as a pilot project from the start

and so the first steps were given on choosing the technology to be used and deciding the scope of the sample. Finally, the pilot implementation was carried out during the 2022-23 academic period, September 2022 and June 2023.

3.2. Participants

It was decided to downsize the scope of the sample to get a manageable representative group of students doing the two most demanded languages in our school: English and Basque. They were selected by the EOIDNA direction board from the two largest groups of students: Secondary Education and the *That's English program*.

The EOIDNA got in contact with the IES (*Instituto de Enseñanza Secundaria*) Valle del Ebro from Tudela and the IES Amazabal BHI from Leitza, who finally volunteered to undergo the pilot experience with English B2 (N=45) and C1 (N=52) level tests respectively. In turn, students doing the That's English Exam Prep B2 and C1 online course were to sit the tests digitally as well. Table 1 illustrates the sample composition:

Table 1. *Comparative Test Results (Source: EOIDNA records)*

Origin	Number of students	Digital language test
IES Valle del Ebro (Tudela, Navarra)	45	English B2
IES Amazabal BHI (Leitza, Navarra)	52	Basque C1
That´s English Exam Prep Online course	44	English B2 & C1

3.3. Instruments and Data Analysis

Several instruments have been used to carry out this study. The tests were digitalised and administrated by the EOIDNA teaching staff with Trelson Assessment software developed on Google (https://www.trelson.com/assessment/) in the case of Secondary students. Teachers and students use their Google institutional accounts to log in. Teachers can follow how the students progress in real time, and the student can continue working even if the network connection is lost. Everything is saved on Drive, even if the computer crashes or the battery runs out of power.

In the case of the That's English Prep Online course students, the tests were digitalised and administrated through the Moodle platform, where the whole course had been lodged. A group of these students opted for doing the test in a face-to-face invigilated room (n=27) whereas another group (n=17) did it from their homes through proctoring software under the supervision of a teacher and a technician. Access was only allowed through the official account of the course. Simulation or mock sessions were organized in all the cases to familiarize the students with the procedure beforehand.

Two additional data sources were used to analyze the influence of digital assessment on student academic performance and students' attitudes or perceptions of the new modality: the contrast of academic performance indicators (success or achievement rates, and average marks obtained) between the digital examinations and the print-based ones in the same call (June 2023); and a questionnaire administered to a sample of students (n=144) on their perception of the online assessment experience (See appendix).

4. Results

4.1. Institutional Preliminary Action

The application of this project has required several previous formal steps: its initial approval, the recruitment of students and the implementation of a simulation or preliminary test. The education authorities showed their support from the start with reasonable reservations that can be expected regarding the willingness of students to sit the tests digitally and how their performance might be affected. The quality of the tests had to be preserved and the project should also have the approval of the pilot students, for whom the test had a definitive value in obtaining the corresponding linguistic certificate.

Therefore, a manageable and representative sample of the two most represented languages in the school, English and Basque, was chosen. Students were also recruited from the two largest groups of students: Secondary Education and the That's English program. The project was proposed to the heads of the selected Secondary Institutes, who agreed to participate counting on the students' consent. In the case of the *That's English program*, students doing the Exam Prep Online course were decided to be the target group.

EOIDNA organized a simulation or preliminary test a month before the real test to familiarize students and detect possible malfunctions. Secondary school students held all the tests face-to-face whereas the online course students did them online too, thus replicating real test conditions as far as possible. Some preestablished aspects were analyzed: password management to access, moving along the different sections of the test, screen visualization, completing and submitting answers by students, and identity checking in the case of proctored students. Those sessions ran smoothly and no major problems were encountered, which initially made the whole project feasible.

4.2. Technological Requirements

The technical aspects faced in our pilot experience have been the following: the choice of the necessary software for test design, having the necessary computer equipment to carry out the tests, relying on an efficient system to register and store students' test answers and allowing tests distribution among teachers for remote correction. We will review each of these parameters in the following sections.

4.2.1. Software Selection and Test Design

The intended software had to allow the administration of the EOI tests with the maximum guarantee of quality. This is, making digital assessment authentic, accessible, appropriately automated, continuous and secure, quoting formerly referred research (Pauli and Ferrell, 2020). Trelson assessment software was chosen since it offers a familiar user interface, making it very easy for both teachers and students to get started and move along its main constituents. EOI tests have a defined, simple and particular structure at the same time which students are used to and expect to find. Thus, the chosen software allows replicating traditional question formats to ensure the authenticity of the tests.

Access is simplified since students do the tests on their computers and use their institutional accounts to log in. With its text-to-voice feature, this software is usable by those who have a long-term auditory or visual disability or a short-term injury, aside from the use of screen readers to help the visually impaired, thus minimizing cost and additional adaptations for particular needs.

Trelson assessment software allows the proper automation of all the stages involved in EOI testing. It eases teachers' test design, marking and feedback workload and can provide quick detailed feedback for students. Even though EOI tests serve as an end-of-course summative examination to validate students' knowledge and language competence, this software also may be used to apply tests as practice opportunities for assessment. In this sense, the tool can also be used as formative continuous assessment teaching students to learn.

Trelson assessment also provides a secure test environment by keeping students focused on their tasks. Additional features include setting a tight testing time limit, randomizing questions and options, and withholding answers until the exam is completed by all students, among others, thus reducing the likelihood of cheating as well.

4.2.2. Equipment and Test Administration

As aforementioned, students answered the face-to-face written and oral tests using their devices in locked mode without the possibility to communicate with other devices, and while invigilated by teaching staff in the room. They used Chromebook devices financed by the Government of Navarra and intended for students of the corresponding Secondary Institutes for their individual use. This enabled centralized control in some key issues like the installation of the Kiosk application, necessary to guarantee security throughout the process. Apart from the odd incidents to be expected, such as having a low battery available in the device or exiting the planned sequence due to an unintentional error, no other significant problems occurred during the tests and all the students could complete them within the established time limits.

In the case of remotely proctored tests using Moodle, students used their equipment once installed the pertinent software under the supervision of teaching staff through a Zoom meeting. The biggest challenge implied checking the students' identities before beginning the tests and the students forgetting login credentials or unintentional disconnecting which were reasonably quickly solved by the supervising team. Having applied a simulation or mock test with some advance hugely contributed to the normal evolution of the examination. It must also be noted the absence of connection drop-out throughout the testing time.

4.2.3. Students' Test Storage and Marking

The files containing students' answers are saved and automatically stored in Google Drive making constant access to them very quick. Another feature of this software permits the manager to upload a group of correctors and share the desired files with each of them so that they can mark the open-question tests. In turn, multiple-option questions are automatically graded. In both cases, Trelson and Moodle, it is possible to quickly access the results and compile statistics later.

4.3. Academic Performance and Student Perceptions

4.3.1. Success and Achievement Rates

To study the performance of the students in the tests, we will analyze the average grades of each group of students in each skill and the percentage of final passes in the call. In this way, we will be able to compare them with the results obtained by the rest of the students who took the paper test in the traditional way. Table 2 shows the results broken down.

Table 2. *Comparative Test Results (Source: EOIDNA records)*

Test	Modality	Number of students	Average grades per skill (Over 20 marks)*					Pass (N)	Pass (%)	Fail (N)	Fail (%)
			CTE	CTO	PCTE	PCTO	M				
English B2	Digital & Face-to-face Moodle	13	4.29	3.78	0.34	12.00	7.68	3	23.1	10	76.9
	Digital & Proctored Moodle	9	11.45	14.18	11.27	9.25	9.13	2	22.2	7	77.8
	Digital & Face-to-face Trelson	45	14.00	16.80	12.40	12.40	11.60	29	64.4	16 (	35.6
	Print-based	1027	14.05	14.30	11.26	11.02	11.83	349	31	772	69
English C1	Digital & Face-to-face Moodle	14	14.78	15.33	11.09	9.29	10.21	5	35.7	9	64.3
	Digital & Proctored Moodle	8	15.00	16.54	10.88	9,70	8.40	2	25	6	75
	Print-based	143	14.25	14.33	10.00	9.27	8.45	40	28	103	72
Basque C1	Digital & Face-to-face Trelson	52	15.37	16.72	13.11	15.60	13.52	49	94.2	3	5.8
	Print-based	1109	15.07	15.41	11.55	11.11	11.41	446	39	705	61

* CTE (Comprensión de textos escritos, Reading Comprehension); CTO (Comprensión de textos orales, Listening Comprehension); PCTE (Producción y coproducción de textos escritos, Writing); PCTO (Producción y coproducción de textos orales, Speaking); M (Mediación, Mediation)

Some trends can be observed before delving into further analysis. In general terms, most students fail the level in the print-based tests in all groups. It must be noted that in order to get a final pass, students need to pass each language area or skill with at least 50% of the score (10 out of 20 marks) and then get 65% of the total score of the test (65 out of 100 marks, once added all the five language areas or skills), as stated by the aforementioned Royal Decree 1/2019, which establishes the common basic principles of evaluation to be applied to the official certification tests of B1, B2, C1 and C2 levels.

In the same way, CTE and CTO offer better results than other skills in all groups. PCTO and M are the only skills with negative scores under 10 marks out of 20 in some groups. In Basque language exams results seem extraordinarily good since this is an official language in Navarre and practically the totality of test takers have been doing their Primary and Secondary studies in the language for years, apart from it being the mother tongue of the students doing the digital test in IES Amazabal from Leitza (PCTO score is 15.60 as compared to 11.11 from the rest of students in C1 level).

There is also a correlation between the pass and fail percentages of all groups of students, except for the groups doing the test digitally with Trelson, who perform far better reversing the trend. The preliminary selection of schools done by the EOIDNA to carry out this pilot experience may account for this fact. IES Amazabal and IES Valle del Ebro, both including around 50 students each, had got good results in previous calls and this would pave the way for a smooth transition into the digitalisation of tests. Thus, the pilot study could mainly focus on the procedure itself.

A more detailed analysis shows similarities and differences between print-based and digital tests, as can be seen in table 3.

Table 3. *Comparative Results Between Print-Based and Digital Tests (Source: EOIDNA records)*

Test	Modality		Number of students	Average grades per skill (Over 20 marks) CTE	CTO	PCTE	PCTO	M
English B2	Digital & Trelson	Face-to-face	45	14.00	**16.80**	**12.40**	**12.40**	11.60
	Print-based		1027	14.05	14.30	11.26	11.02	11.83
Basque C1	Digital & Trelson	Face-to-face	52	15.37	**16.72**	**13.11**	**15.60**	13.52
	Print-based		1109	15.07	15.41	11.55	11.11	11.41

Students taking the digital face-to-face test got higher mean scores in CTO, PCTE and PCTO language areas in both language tests (highlighted in bold) and similar ones in the rest. In the listening and writing tests, students controlled the audio file, listened with earphones and edited their writings on the computer under time limits, which might have affected it. As for the speaking tests, students could see the questions on their screens permanently, which might have also had them more focused on the task.

If compared within the same digital modality, whether under direct classroom supervision or proctored one, there seem to be smaller differences in general as shown in table 4. The reduced size of the sample must always be taken into account.

Table 4. *Comparative Results Between Digital Face-to-Face and Proctored Tests (Source: EOIDNA records)*

Test	Modality	Number of students	Average grades per skill (Over 20 marks) CTE	CTO	PCTE	PCTO	M	Pass (N)	Pass (%)	Fail (N)	Fail (%)
English B2	Digital & Face-to-face Moodle	13	14.29	13.78	10.34	12.00	7.68	3	23.1	10	76.9
	Digital & Proctored Moodle	9	11.45	14.18	11.27	9.25	9.13	2	22.2	7	77.8
English C1	Digital & Face-to-face Moodle	14	14.78	15.33	11.09	9.29	10.21	5	35.7	9	64.3
	Digital & Proctored Moodle	8	15.00	16.54	10.88	9,70	8.40	2	25	6	75

In this case, evidence indicates that students did not get higher scores on online proctored tests than on face-to-face ones as a general rule.

4.3.2. Student Perceptions

After the pilot experience, students were asked about having undertaken tests digitally through a questionnaire included in Appendix 1. In this sense, we will analyze several aspects below.

A sample of 69 responses was obtained (49% of the students who took the test, a total of 141). 60 respondents (87%) did the test in a supervised face-to-face classroom, whereas 9 (13%) opted for a remote digital test through proctored software. 60 respondents (87%) had previously carried out a voluntary drill or training mock test. 31 respondents (44.9%) took the Basque test, while 38 (55.1%) took the English test. 29 of them (42%) did the B2 level test and 40 (58%) did the C1 test.

Regarding the emotional effect that taking the test digitally had on them, 31 (44.8%) recognized a positive or encouraging influence, 27 (39.1%) manifested themselves in a neutral way and 11 (15.9%) admitted a somehow negative influence. When asked about possible problems suffered during the exam, 34 students (49.3%) stated that they had not suffered any, 33 (47.8%) said they had had a small incident that was quickly resolved, while 2 (2.9%) reflected having had more serious incidents.

Respondents rated the previous drill practice as satisfactory (3.5/5 on a Lickert scale, 1 being very unsatisfactory and 5 very satisfactory). In the same way, they valued the visualization of the text on the screen at 3.62 and the sound quality of the audio at 3.96. The possibility of completing one part of the test and being able to move towards the next one individually with flexibility was valued positively at 4.16/5. Finally, they valued their degree of satisfaction with the digital test in general at 3.88/5. The time allowed for the test also seemed convenient to them (3.94/5).

37 students (53.6%) considered that the test did not seem easier or more difficult for having taken it digitally. 22 respondents (31.9%) believed it was easier while 10 (14.4%) considered it more difficult. If asked which exam method they prefer after this experience, 27 respondents (39.1%) preferred the digital face-to-face exam with their device in a supervised classroom, 16 (23.2%) preferred to do it on paper, 14 (20, 3%) digitally from their location, and 12 (17.4%) considered that it is indifferent.

5. Discussion

This study aimed to analyze the implementation of a pilot digitalised test that replaces a print-based one to check the challenges and opportunities this process entails for EOIDNA as a public educational institution in an ordinary examination call.

After the experience, it can be stated that this transformational process has required undertaking some key formal steps to prove feasible and run smoothly: a thoughtful plan and its initial approval, the prudent recruitment of pilot students and the implementation of an authentic preliminary test. The agents involved in the process, educational authorities, teaching staff, students and technology developers have needed to join efforts to guarantee its viability and solidity. In our case, digital assessment has proved a convenient and effective tool to maximize the benefits for all involved stakeholders (Timmis, et al., 2015; Laborda et al., 2015; Okada, 2019) through a quickly accessible test, its flexible administration, and an efficient management and storage of student test records and grades.

The technology used, Trelson software and Moodle, has allowed EOIDNA to preserve the authenticity of traditional EOI tests. The sections of the tests have been replicated and all the administration requirements have been applied to the letter. Truth be told, by operating on Google, Trelson software has outdone Moodle as a simpler and integrative tool that comprises test design, test distribution among remote correctors and operational agility within the same tool. Both tools proved easy to use for students and no major technical problems were found in the administration of the test, as results have shown. So far, digital assessment has managed to become authentic, accessible, appropriately automated, continuous and secure enough (Pauli and Ferrell, 2020). A future challenge to face may remain in making the test scalable and extensive to a far more numerous group of test candidates.

In our study, students' academic performance followed similar general trends in both digital and print-based test modalities, as was suggested by previous research (Butler-Henderson and Crawford (2020). In our view, this finding objectively reinforces the validity of digital tests and may help reduce initial prejudices against them. Doubts about the integrity of online tests may also lead to the assumption that students can cheat in exams, and as a result, their performance would be higher. However, evidence in our study indicates that proctored students did not outperform face-to-face ones on online tests (Dominguez-Figaredo et al., 2022; Alessio et al., 2018; Arnold, 2016; Hylton et al., 2016; Cheng et al., 2016).

When it comes to the challenges that digital tests entail for students, a vast majority (84%) reported a positive or neutral effect on them the very first time they took the whole test on their computers. They rated the experience as highly satisfactory and especially appreciated the quality of the audio and the flexibility of the test for being able to move through it at their own pace. Around 60% of them showed their preference for an online test (either face-to-face or proctored) and an additional 17% would not mind doing it that way in the future as they thought it would make no difference.

This study suffered from several shortcomings which limit the generalizability of the findings. First, it was not based on many participants, nor were the participants randomly selected. Future research should help overcome the limitations of this study mainly by expanding the size of the sample of each type of test, though the fact that students are taking real value tests should never be understated.

6. Conclusion

To conclude, our findings seem reasonably sound and stimulating for our institution. In the future, some regulatory support from the education administration will also be welcome to advance towards widespread digital assessment methods which can be replicated and eventually implemented in other EOI. We know we need to be very cautious about moving forward, with enough time to plan and evaluate the impact of the changes. We firmly believe that technology-mediated assessment needs to be implemented out of conviction, as has been our main aim in this study.

Acknowledgements

We would like to express our gratitude to the students and teaching staff involved whose engagement and generous availability have made this pilot study possible.

Conflict of Interest

Not applicable.

References

Alessio, H. M., Malay, N., Maurer, K., Bailer, A.J., & Rubin, B. (2018). Interaction of proctoring and student major on online test performance. *International Review of Research in Open and Distributed Learning*, *19*(5), 166–185. https://doi.org/10.19173/irrodl.v19i5.3698.

Arnold, I. J., 2016. Cheating at online formative tests: Does it pay off? *The Internet and Higher Education, 29*, 98-106. https://doi.org/10.1016/j.iheduc.2016.02.001.

Arrosagaray, M., González-Peiteado, M., Pino-Juste, M., & Rodríguez-López, B. (2019). A comparative study of Spanish adult students' attitudes to ICT in classroom, blended and distance language learning modes. *Computers & Education, 134,* 31–40. https://doi.org/10.1016/j.compedu.2019.01.016.

Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education, 159,* 104024. https://doi.org/10.1016/j.compedu.2020.104024.

Cheng, L. K., Puteh, F., Selamat, A., & Mohamed, F. (2016). A review of recent methodologies in English language content delivery. In *2015 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)* (pp. 169-174). IEEE. https://doi.org/10.1109/IC3e.2015.7403507.

Consejo Escolar del Estado. (2022). *Informe 2022 sobre el estado del sistema educativo. Curso 2020-2021*. Accessed 12 June 2023 at https://www.educacionyfp.gob.es/mc/cee/publicaciones/informes-del-sistema-educativo/informe-2022.html

Darabi Bazvand, A., & Ahmadi, A. (2020). Interpreting the validity of a high-stakes test in light of the argument-based framework: Implications for test improvement. *Journal of Research in Applied Linguistics*, *11*(1), 66-88.

Domínguez-Figaredo, D., Gil-Jaurena, I., & Morentin-Encina, J. (2022). The impact of rapid adoption of online assessment on students' performance and perceptions: Evidence from a Distance Learning University. *The Electronic Journal of e-Learning, 20(3),* 224-241. http://doi.10.34190/ejel.20.3.2399.

European Commission, European Education and Culture Executive Agency, Birch, P., Baïdak, N., De Coster, I. et al. (2023). Key data on teaching languages at school in Europe. *Publications Office of the European Union,* https://data.europa.eu/doi/10.2797/529032.

Gold, S. S., & Mozes-Carmel, A. (2009). A comparison of online vs. proctored final exams in online classes. *Journal of Educational Technology, 6*(1), 76–81. https://doi. org/10.26634/jet.6.1.212.

Hernández, A., Elosua, P., Fernández-Hermida, J. R., & Muñiz, J. (2022). Comisión de test: Veinticinco años velando por la calidad de los test. Test commission: Twenty-five years working on test quality. *Papeles del Psicólogo / Psychologist Papers, 43*(1), 55-62. https://doi.org/10.23923/pap.psicol.2978

Hylton, K., Levy, Y., & Dringus, L.P. (2016). Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education, 92*, 53-63. https://doi.org/10.1016/j.compedu.2015.10.002.

Inbar-Lourie, O. (2013). The special issue on language assessment literacy. *Language Testing, 30*(3), 301-307.

Laborda, J. G., Sampson, D. G., Hambleton, R. K., & Guzman, E. (2015). Guest editorial: Technology supported assessment in formal and informal learning. *Educational Technology & Society, 18* (2), 1–2.

Makhachashvili, R., & Semenist, I. (2022). Systemic dynamics of digital education in foreign languages programs. In *Proceedings of the 16th International Multi-Conference on Society, Cybernetics and Informatics: IMSCI 2022*, (Vol. 1, pp. 5-10). International Institute of Informatics and Systemics, USA. https://doi.org/10.54808/IMSCI2022.01.5

Nurieva, G. R., & Garaeva, L. M. (2020). Zoom-based distance learning of English as a foreign language. Special Issue Proceedings of the 7th International Conference on Applied Linguistics Issues (ALI 2020). *Journal of Research in Applied Linguistics, 11*(3), 439-448.

OECD. (2020). Policy responses to COVID-19. Accessed 25 June 2023 at https://www.oecd.org/coronavirus/en/policy-responses.

Okada, A., Noguera, I., Alexieva, L., Rozeva, A., Kocdar, S., Brouns, F., Ladonlahti, T., Whitelock, D., & Guerrero-Roldán, A. E. (2019). Pedagogical approaches for e-assessment with authentication and authorship verification in Higher Education. *British Journal of Educational Technology, 50*(6), 3264–3282. https://doi.org/10.1111/bjet.12733

Öz, H., & Özturan, T. (2018). Computer-based and paper-based testing: Does the test administration mode influence the reliability and validity of achievement tests? *Journal of Language and Linguistic Studies, 14*(1), 67-85.

Pauli, M., & Ferrell, G. (2020). The future of assessment: five principles, five targets for 2025. JISC. Available at: http://repository.jisc.ac.uk/7733/1/the-future-of-assessment-report.pdf. [Accessed 28 August 2023]

Razavipour, K. (2013). Assessing assessment literacy: Insights from a high-stakes test. *Journal of Research in Applied Linguistics, 4*(1), 111–131.

Soodmand Afshar, H. (2020). Test-takers' perceptions of paired speaking tests and the role of interlocutor variables in pairing. *Journal of Research in Applied Linguistics, 11*(1), 89-123.

Stickler, U., & Shi, L. (2016). Tell us about CALL: An Introduction to the Virtual Special Issue (VSI) on the Development of Technology-Enhanced and Computer-Assisted Language Learning published in the System Journal. *System, 56,* 119–126. https://doi.org/10.1016/j.system.2015.12.004

Suárez-Álvarez, J., Fernández-Alonso, R., García-Crespo, F. J., & Muñiz, J. (2022). The use of new technologies in educational assessments: Reading in a digital world. *Papeles del Psicólogo / Psychologist Papers, 43*(1), 36-47. https://doi.org/10.23923/pap.psicol.2986

Timmis, S., Broadfoot, P., Sutherland, R. & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal, 42*(3), 454–476. https://doi.org/10.1002/berj.3215

Appendix

EOIDNA DIGITALISED ASSESSMENT: STUDENT SATISFACTION QUESTIONNAIRE

Personal information

1. I have taken the digital test...
   a) In a supervised classroom
   b) From my location (Home, work, etc.)
2. I have previously carried out a drill to familiarize myself with the digitized test...
   a) Yes
   b) No
3. Language test that I have taken...
   a) Basque
   b) English
4. Language level of the digital test that I have taken...
   a) B2
   b) C1

Assess the following aspects about the digitalised test

5. Having to take the test on a computer...

a) It has negatively influenced me
b) It has discouraged me somehow
c) It has neither encouraged nor discouraged me
d) It cheered me up somehow
e) It has had a positive influence on me

6. During the digitized test...
   a) I have suffered serious incidents
   b) I have suffered some minor incident
   c) I have not suffered any incident
7. My previous familiarization with the digital test through a simulation has been…
   Rate between Unsatisfactory (1) - Satisfactory (5)
8. The display of written texts on the screen...
   Rate between Deficient (1) - Excellent (5)
9. The quality of sound in audio files...
   Rate between Deficient (1) - Excellent (5)
10. The ability to complete a skill and advance to the next individually…
    Rate between Inconvenient (1) - Convenient (5)
11. My degree of satisfaction with the digital test in general…
    Rate between Unsatisfactory (1) - Satisfactory (5)
12. The time to take the digital test seemed…
    Rate between Scarce (1) - More than enough (5)
13. I consider that the digitized test is…
    a) Far more difficult than the paper test
    b) More difficult than the paper test
    c) Neither easier nor more difficult than the paper test
    d) Easier than the paper test
    e) A lot easier than the paper test
14. After taking the digital test, which modality do you prefer for future tests?
    a) Paper tests
    b) Digital tests with a computer in a supervised classroom
    c) Digital tests with a computer from my location
    d) It do not mind