

Students' Opinions about Ubiquitous Delivery of Standardized English Exams

LITZLER, MARY F.

JESÚS GARCÍA LABORDA

Universidad de Alcalá

Received: 2 September 2015 / Accepted: 1 January 2016

ISSN: 1697-7467

ABSTRACT: This paper discusses the results of a study conducted with 218 students in the final year of high school to determine their opinions about the feasibility of using a tablet PC for delivery of a standardized English language test. One such test could be the English paper of the exam given to students upon completion of the Baccalaureate program in Spain. The results of the questionnaire used reveal a positive tendency towards the platform and application, particularly regarding the organization and hierarchy of the test and to some extent its intuitive nature. Nevertheless, a high level of indecision was observed.

Keywords: norm-referenced evaluation, computer-assisted examination, interface, survey report, tablet PCs.

Opiniones de los estudiantes sobre la realización ubicua de exámenes estandarizados de inglés

RESUMEN: Este artículo trata sobre los resultados de un estudio llevado a cabo con 218 alumnos en su último curso de escuela secundaria para determinar sus opiniones sobre la posibilidad de utilizar tabletas informáticas para la realización de un examen estandarizado de inglés. Los resultados del cuestionario revelan una tendencia positiva hacia el uso de la plataforma y la aplicación, en particular en lo referente a la organización y jerarquía del examen y hasta cierto punto la naturaleza intuitiva del examen. Sin embargo, se observó un alto grado de indecisión hacia los efectos previsibles del uso de este medio de realización.

Palabras clave: evaluación normativa, examen asistido por ordenador, interfaz, informe sobre encuesta, tabletas.

1. INTRODUCTION

Modern technology has opened up many opportunities for learning. Since the development of the Internet, teachers and their students have been able to work remotely using e-mail, virtual learning platforms such as Blackboard, wikis, and more recently, social media. The development of modern technology is also of use in the area of testing, even though this area has lagged somewhat behind compared to teaching and learning, especially in Spain (García-Labordá, 2010; García-Labordá, Magal-Royo, Litzler & Giménez López, 2014).

Some standardized testing institutions such as Educational Testing Service (TOEFL and TOEIC), the British Council (APTIS), Cambridge English Language Assessment (KET, FCE etc.) and IELTS have already made the transition to technology for delivering exams. There are several advantages to using modern technology for evaluation. It motivates the students

(Sapriati & Zuhairi, 2010), enables fast correction of standardized exercises (McNulty, 2011), allows test correctors to supervise the corrections while at a different location (García-Laborda, 2007), even in a different country, and it may prove to be less expensive than other delivery formats (Bulut & Kan, 2012). The use of tablet PCs in particular for delivery of standardized tests also has some potential benefits. First, they combine the advantages of desktop PCs and the Internet but occupy a smaller amount of space and are easily stored and transported (Gawełek, Spataro & Komarny, 2011). For this same reason, they can be used by more than one center if exam dates do not coincide. At the same time, thanks to their cheaper prices as compared to desktop PCs, tablet PCs are a familiar household item for many young people today, and this means that students could adapt to them easily for educational purposes.

Recent legislation in Spain (LOMCE, 2013) calls for high stakes English exams in the 4th, 6th, 10th and 12th years of school. The purpose of the exams is to serve as a completion to the different cycles of education but, in the case of the 10th and 12th grades, they can have an impact on the students' possibilities for attending the two-year university preparation course and the university itself, respectively. At the same time, an oral component is planned to be added to the official test (formerly the university entrance exam) taken at the end of the Baccalaureate program in this country. These additional testing situations will lead to an increased demand on educators' time. For this reason, new formats of exams that can facilitate the testing of students and the correction of the tests are welcome. Tablet PCs and other ubiquitous delivery means are one possibility of innovation in this area (García-Laborda et al. 2014).

The OPENPAU project, which has been financed by the Ministry of Education and the Ministry of Science, has been examining the potential use of tablet PCs for foreign language testing, in particular English, in high schools. Different phases of the project have examined teachers' attitudes towards the use of tablet PCs for testing, ergonomics, usability, reliability, design and interface factors in their use (García-Laborda et al., 2014). The present study examines the reactions of a group of students to the feasibility of using tablet PCs and desktop computers for English testing. After participating in a mock exam using these two delivery formats the students indicated their opinions in a questionnaire. The results obtained for the use of the tablet PCs are reported here. This research provides the first results reported in terms of student attitudes towards the use of tablet PCs for standardized testing of English as a foreign language in Spain.

1.1. Technical specifications of the OPENPAU platform

At present a number of open source applications are available for the ad hoc creation of teaching and learning materials which can be shared online through educational communities around the world (García-Laborda, 2009). One widely used application of this sort is Moodle©. It allows for the creation of new applications for learning as well as generic modules for evaluation through multiple choice questions, short answer questions, etc., all of which can be imported from other systems and exported to them (Gimenez-Lopez, Magal-Royo, García-Laborda, & Garde-Calvo, 2009). Nevertheless, it does not have a specific module for evaluating linguistic competence in a foreign language. For this reason, the OPENPAU project is using the Android SDK developer module to create a highly specialized application compatible with Moodle©. The final result will be appli-

cable to any Moodle© platform used by any educational institution and it will enable analysis of both qualitative and quantitative information obtained during foreign language proficiency testing to be studied by teachers and administrators. In developing the application, the OPENPAU technical team is focusing on three aspects considered fundamental for the validation of the end version: compatibility among systems, multimodality of the environment, and data security.

The programming language used is Moodle© PHP and the application can run on Linux©, BSD©, Mac OS-X© and Windows©. For the development of the application Linux© server with Apache© MySQL 5.3.2+ 5.0.25+ server are being used. The assembly is being developed in XHTML + CSS so that it can be properly displayed on mobile devices such as tablet PCs. The tablet PC used for this research was a Wolder miTab EVOLUTION W2 (figure 1) 10.1" HD IPS reinforced, with QUAD CORE, 16 GB ,and a QWERTY BT keyboard.

Figure 1. Research materials (desktop PC, tablet PC, headphones).



2. METHOD

This study required the student participants to complete a mock English test delivered through a trial version of the application being designed by the OPENPAU technical team. The data for this paper were collected through a questionnaire. The mock test and questionnaire are described in more detail in Section 2.2.

2.1 Participants

A total of 218 students in the second year of the Baccalaureate program in high school took part in this study. Visits were made to 9 schools in the northeast part of the region of Madrid and the province of Guadalajara, a random sample which represents about 9% of the estimated population of the area covered.

Table 1. Schools and participants involved in the study.

	Total	Percentage
IES Brianda de Mendoza	15	6,9
IES Aguas Vivas	22	10,1
IES Clara Campoamor Yunquera	30	13,8
IES Valle Inclán	41	18,8
IES Emperatriz María de Austria	58	26,6
IES Alcal'a Nahar	23	10,6
IES Albeniz	10	4,6
IES Antonio Machado	12	5,5
IES Cardenal Cisneros	7	3,2
Total	218	100,0

In fact, the schools correspond to three socio-economic categories: low, middle and upper-middle class, providing reason for the acceptability of the sample despite some limitations mentioned below.

2.2. Procedure and instruments

2.2.1. Testing sessions and mock test

The study was conducted between February and March 2015. In order to perform the research, the OPENPAU team obtained permission from the school principals and the students, as most of them were 18 years or older and could sign a written agreement. The sessions were conducted during the students' English classes or during free periods in the morning and they took place in the school computer laboratories for a span of two hours each. In the first part of each session, the researchers and technicians explained the testing platform and provided instructions on how to complete the mock exam exercises described below. The students then completed the exercises with the tablet PCs provided by the research team (few were available at the schools) and afterwards they did the same exercises

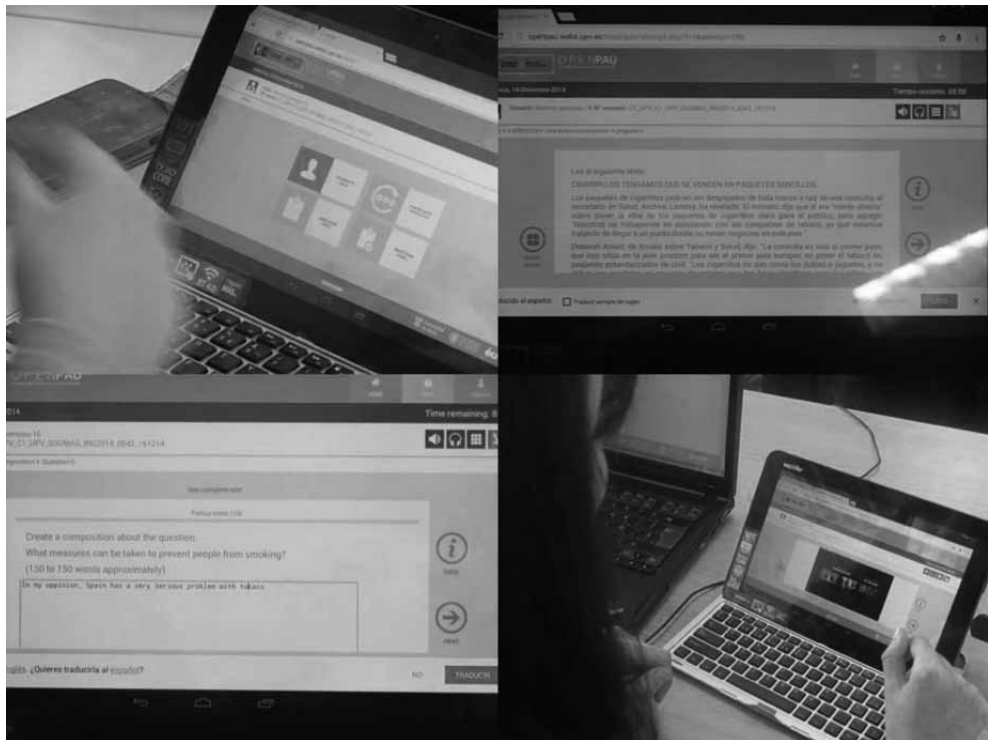
using the desktop PCs in the computer laboratory. In the final part of each testing session, the participants filled out the questionnaire described in the next section.

It must be recalled at this point that the objective of the study was to determine the students' responses to the two delivery formats, as opposed to their actual results on the exercises. For this reason, it was decided that the exercises completed in the two delivery modes should be identical. If the exercises had been different, the students might have been influenced by the difficulty of one set of exercises over the other in indicating their opinions on the questionnaire for one of the two delivery formats.

The mock exam contained questions that examined proficiency in reading, writing, listening and speaking. There were five types of questions (Figure 2):

- 1) 2 open questions in response to a 250-word reading
- 2) 2 True/False questions about the same reading
- 3) 4 multiple choice questions on grammar topics
- 4) 1 150-word essay about one of two possibilities
- 5) 3 open speaking questions in response to a 1-minute video

Figure 2. Different sections of the test (task selection, reading, writing, speaking).



Throughout the test, the participants could use a pen and paper to prepare for the essay or take notes while the video was played. Upon completion of the mock exam on the tablet PC and the desktop PC, they needed to validate their responses to finish the test. At this point, the platform redirected them to the online questionnaire.

2.2.2. *Questionnaire*

The questionnaire was developed by a team of teachers, linguists and computer technicians using the Dephi method (Williams, Boone & Kingsley, 2004; Rice, 2009; Hsu, Ching & Snelson, 2014). According to this process, the initial questions were written by the experts and then piloted with a first group of students. At this point, they were reviewed by the experts and modified before they were piloted with a second group of students. This second piloting of the research instrument was conducted with a group of 34 pre-service teacher trainees (García-Laborda, Magal-Royo, Rodríguez-Lázaro & Fuentes-Marugán (2015)). Afterwards, the questionnaire was used with the entire population for this study.

The topics of the questions covered test design, interface design, visual ergonomics, student reception and others, as can be seen in Table 2 below, and they followed a 5-point Likert scale varying from “strongly disagree” to “strongly agree.” Nine of the questions were stated in the affirmative but question 10 was formulated in the negative. The reason for including this latter form was to avoid an acquiescent bias in which the students would agree with all of the questions without really thinking about their responses (Fischer, 2003). In addition, questions 4 and 8 were identical and served as a control to determine whether the participants were attentive in answering the questions, particularly given the length of the research sessions (two hours). The questionnaire was delivered in Spanish, the native language of the respondents.

After the 218 questionnaires had been completed, the responses to each of the questions were tallied and the percentages for each of the options selected on the 5-point Likert scale were calculated. The scoring was done in the traditional weak to strong agreement, so a value of one represented the most negative category and five the most positive. Chi-square analyses were then calculated using the SPSS 16 software to determine whether the results obtained for each of the questions were significant. Only one of the results was found to be significant, as indicated below.

3. QUESTIONNAIRE DATA RESULTS

The return rate of the questionnaires came to 91.9% of the survey population (218 participants) because the students completed it on the desktop computer during the testing session. The remaining 8.1% of uncompleted questionnaires (19 students) corresponded to students who did not have time to complete the survey after doing the mock test and, therefore, left the classroom. This last group of students were not included in the number of participants indicated in Section 2.1.

The most striking aspect of the results obtained in this research was the high level of indecision as indicated by the relatively large numbers of students who chose the middle

option on the Likert scale for the questions. In five of the ten questions, more than 30% of the participants opted for this value, reducing the overall responses to the negative and positive values. This tendency may have been an important factor in the chi-square calculations, which revealed that the results to the different questions were generally not statistically significant. Nevertheless, when this high degree of indecision is ignored and the remaining percentages are examined, a tendency towards a positive view of the tablet PC exam format is observed in the individual question results as well as in the overall percentages for attitude. Specifically, the average percentages of positive responses (values 4 and 5) and negative responses (values 1-2) of the affirmative statements (questions 1-9), come to 44% and 26% respectively. In this sense, the findings of this study revealed that the students generally thought that the platform and the application were adequate for delivery of a standardized English test.

Table 2. Values obtained in the questionnaire.

	1	2	3	4	5
	Percentages (%)				
	Number of responses				
1) The visual organization seems adequate to me for a student in the 2nd year of Baccalaureate studies.	8,3 18	8,3 18	22,5 49	25,2 55	35,8 78
2) The interface elements allow me to move around the application easily.	12,8 28	15,6 34	31,7 69	27,1 59	12,8 28
3) The interface is very attractive.	13,3 29	19,3 42	35,8 78	23,4 51	8,3 18
4) The order and hierarchy of the exam are clear.	10,1 22	12,8 28	23,9 52	24,8 54	28,4 62
5) The icons and graphic symbols enable me to do the exam easily.	11,9 26	7,8 17	38,1 83	25,2 55	17,0 37
6) The application seems adequate to me for taking this exam.	21,1 46	11,5 25	27,5 60	21,6 47	18,3 40
7) The application seems intuitive to me.	12,8 28	8,3 18	36,7 80	29,4 64	12,8 28
	1	2	3	4	5

	Percentages (%)				
	Number of responses				
8) The order and hierarchy of the exam are clear.	11,9	11,5	28,9	27,1	20,6
	26	25	63	59	45
9) A student in the 6th year of primary school would have problems doing an exam on this platform.	14,7	17,4	30,3	20,6	17,0
	32	38	66	45	37
10) Some things do not work the way I think they should.	11,0	13,3	24,8	21,6	29,4
	24	29	54	47	64

Question 1 drew the most positive results of all the questions. Sixty-one percent of the participants agreed or strongly agreed that the visual organization of the tablet PC application was adequate for a student in the final year of high school. It was especially remarkable that the percentage of students who strongly agreed with the statement by selecting the value of 5 (35.8) was the largest percentage obtained for all of the questions, meaning that they considered the visual organization of the application to be excellent.

Question 2 was similar to question 1 but queried the students about a specific aspect of the interface. In this case the results were positive about their ability to move around the application, but less than for the other questions. It was not clear at this point whether the lower positive trend (39.9%) related to interface and orientation in the application was due to a lack of clarity of the visual elements, or whether the interface was considered too simple in a world where interfaces are becoming increasingly richer. Nevertheless, the results were significant as indicated by the Chi-square value (15.3189), where the P value is 0.000091 ($p < 0.05$). The results for question 3 showed a more ambivalent result in terms of the participants' impression of the aesthetics of the interface; the largest percentage of responses corresponded to the middle option on the scale and the negative responses only slightly outweighed the positive ones (32.6% and 31.7% respectively).

Question 4 related to the importance of a clear test delivery design as manifested through the order and hierarchy of the exam. While the students were only somewhat positive towards the organization of the interface, as revealed in the findings for question 2, the results for question 4 indicate that they were positive towards the test organization, with 53.2% of the students selecting options 4 (agree) and 5 (strongly agree). However, the result was not significant ($\chi^2 = 1.2478$; $P = 0.263971$), probably due to the large number of undecided students. Still, it should be considered as a positive indicator the fact that the ratio was nearly 2:1 between students who agreed and those who disagreed regarding the clear organization and hierarchy of the exam.

Question 5 intended to determine whether the visual elements in the interface facilitated completion of the test. In this case, the findings again revealed a large number of positive responses, as 42.2% of the students agreed or strongly agreed with the statement, although the difference was not significant ($\chi^2 = 4.8269$; $P = 0.028019$). This question was loosely connected to question 6, which asked the students if they felt that the applica-

tion was adequate to deliver the test. The responses to this question were also positive (39.9%), though less than in the other questions. In fact, this question drew one of the largest percentages of students who felt negatively about it (32.6%). At the same time, in question 9, a total of 37.6% of the students also indicated that the platform would be challenging for their 6th grade counterparts. Nevertheless, this last question involves a hypothetical situation, which may have caused the participants some difficulty because they had to imagine how the 6th grade students would feel. Despite this limitation, the responses can provide some insight until a population of students of this younger age can be queried directly.

Question 7 was an important one as it considered whether the students felt the platform was intuitive. Technology must not create a gap amongst students in terms of their digital skills and technological know-how when the objective of a test is to determine their ability in English. In this sense, the results were positive since 43.2% agreed or strongly agreed that the platform was intuitive while only 21.1% indicated that it was not intuitive. However, this question drew one of the largest percentages of undecided students (36.7%). Both this degree of indecision and the percentage of students who disagreed with the question could be related to the results of question 10, which was related to expectations about how the application and platform should work. Fifty-one percent of the participants, one of the highest percentages found amongst the data, indicated that there were things about the platform that did not work as they had expected.

Question 8 was a control question that did not reveal any major differences with the responses obtained for question 4. The positive tendency in attitude towards the order and hierarchy of the exam is maintained, albeit to a slightly lesser extent. Specifically, the favorable responses drop from 53.2% in question 4 to 47.7% in question 8, and the negative responses increase from 22.9% to 23.4% respectively.

4. ANALYSIS

Traditionally, there are two main reasons why students and teachers may be reticent about the introduction of new technology: its novelty and the intrinsic difficulty of managing new software (Stockdill & Morehouse, 1992; Hoerup, 2001). It can be extremely stressful for students to take a high-stakes exam (one which can have a major impact on their future) if they have to use a new platform when they are more used to using a pen and paper (Colwell, 2013), or they simply believe that a traditional testing format is more efficient and fair (Marks & Cronje, 2008). The different modes of computer-based testing (desktop PCs, tablet PCs and mobile phones) also involve uncertainty (Bartram, 2006). Some software can create the visual impression of adding difficulties to the test itself (Saade & Kira, 2007); students need to understand how the software works in addition to knowing the content of the exam, which can be difficult without previous training.

At first glance, it could seem that the high percentage of indecision obtained in this study is a manifestation of student reticence to using tablet PCs for high-stakes exams. However, other reasons might be behind this result. First, it may be a consequence of fatigue after doing the tests (Ackerman & Kanfer, 2009), which took more than an hour to complete. Secondly,

it could be linked to the fact that the students are not used to being asked their opinion on official school matters and they did not know what to respond (Rourke & Hartzman, 2008; Wood, 2011). Regardless of the motive behind the high percentage of uncertainty, the students were more positive than negative in their answers to eight out of the nine questions written in the affirmative. In some cases more than 50% of all the responses were in agreement with the statements. In this sense, the results of this study can be considered extremely positive given the existence of reasons to account for student reticence.

In this context, the researchers view the responses to questions 1, 4 and 7, as well as the control question, as being favorable because they indicated that the students had a positive impression of the overall design of the application and platform. The participants considered the application as acceptable for students at their level in high school (the final year) albeit less so for students in the 6th year of primary school (question 9). Test format and design need to be highly intuitive to prevent the introduction of a bias in favor of students who are more technologically experienced. If students are unfamiliar with a platform, they should still be able to guess how it works in order to be able to complete a test (Wise, Pastor & Kong, 2009).

The students were also positive about the visuals, icons and symbols of the application. However, question 3 provided evidence that they did not like the interface. This may be because testing interfaces are dull and simple (Fulcher, 2003), although this simplicity is completely necessary to avoid a bias that can discriminate students based on computer experience instead of on language competence (Fulcher, 2003, García-Laborda, 2007). An elaborate interface might prove distracting to students less accustomed to working with computers.

It is also worthy of note that the students were less clear about whether the platform and application were adequate for the delivery of the test (question 6) and that there were aspects that did not work the way they thought they should (question 10). Neither of these points can be ignored if standardized English tests are to be delivered using tablet PCs and other ubiquitous devices in the future. Nevertheless, the students did show a moderately positive attitude overall towards the design and implementation of a high-stakes English test through this specific platform as can be concluded from the questionnaire results overall.

Despite the positive results obtained here, this study does have a few limitations that deserve attention. A first one is the fact that only 218 students responded to the questionnaire. It is difficult to obtain a larger population sample without the assistance of the official educational institutions, but this research was conducted by an independent team unrelated to the regional school boards. At the same time, the data were obtained only upon completion of the entire research session, so 19 students who participated in the mock exam but did not complete the questionnaire could not be counted in the results. Nevertheless, the findings are of interest, particularly because the sample size is acceptable for research carried out independently. A second limitation is the fact that the data have not been confirmed or clarified to date; this is particularly the case of the large amount of indecision that was observed throughout the questionnaire responses. Focus groups and/or interviews could provide more insight to reasons behind these findings.

5. CONCLUSION

The results of this study are encouraging for the potential of tablet PCs as a means for delivery of standardized exams. Despite a high level of indecision on the part of the students participating in this study, the results obtained with a sample of 218 participants revealed that overall the students felt that a standardized English text could be delivered using tablet PCs. They were particularly positive about the visual organization of the application and the order and hierarchy of the mock exam completed during the research sessions. They also indicated that the platform was intuitive, a fundamental aspect of test format and delivery means if biases in favor of more technologically savvy students are to be avoided.

Despite these positive results, much needs to be done before tablet PCs and other ubiquitous devices can be used for standardized testing of English in Spain. During the OPENPAU project, the researchers have found that many limitations exist in terms of technological development and standards in schools. The different schools contacted as part of the study had completely different technological means available, as there is no standard criteria. Overall, the strength of WI-FI signals varied from school to school and the software and hardware encountered in schools' computer rooms was often starting to become out of date due to the use of older computers using Windows XP. More reliable connections are urgently needed if large-scale testing is to be undertaken.

In addition, in the case of applications for foreign language learning, updated technology is fundamental as students must be encouraged to interact, listen, understand and carry out tasks through audiovisual means in order to simulate a real-world language experience and, hence, increase student motivation. This is a significant reason to update technological capacity in schools. English tests delivered through tablet PCs and other ubiquitous devices are a natural continuation of this shift.

6. REFERENCES

- Ackerman, P. L., & Kanfer, R. (2009). "Test length and cognitive fatigue: An empirical examination of effects on performance and test-taker reactions", in *Journal of Experimental Applied Psychology*, 15, 2: 163-181.
- Bartram, D. (2006). "The internationalization of testing and new models of test delivery on the internet", in *International Journal of Testing*, 6, 2: 121-131.
- Bulut, O., & Kan, A. (2012). "Application of computerized adaptive testing to entrance examination for graduate studies in Turkey", in *Eurasian Journal of Educational Research*, 49: 61-80.
- Colwell, N. M. (2013). "Test anxiety, computer-adaptive testing and the common core", in *Journal of Education and Training Studies*, 1, 2: 50-60.
- Fischer, R. (2004). "Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP", in *Journal of Cross-Cultural Psychology*, 35: 263-282.
- Fulcher, G. (2003). "Interface design in computer-based language testing", in *Language Testing*, 20, 4: 384-408.
- García-Laborda, J. (2007). "On the net: Introducing standardized EFL/ ESL exams", in *Language Learning & Technology*, 11, 2: 3-9.

- García-Laborda, J. (2009). "Interface architecture for testing in foreign language education", in *Procedia Social and Behavioral Sciences*, 1, 1: 2754-2757.
- García-Laborda, J. (2010). "¿Necesitan las universidades españolas una prueba de acceso informatizada? El caso de la definición del constructo y la previsión del efecto en la enseñanza para idiomas extranjeros", in *Revista de orientación y Psicopedagogía*, 21, 1:71-80.
- García-Laborda, J., Magal-Royo, T., Litzler, M.F., & Giménez-López, J.L. (2014). "Mobile phones for a University Entrance Examination language test in Spain", in *Educational Technology & Society*, 17, 2: 17-30.
- García-Laborda, J., Magal-Royo, T., Rodríguez-Lázaro, N., & Fuentes-Marugán, L. (2015). "Facilitating Language Tests Delivery through Tablet PCs", in *Procedia Social and Behavioral Sciences*, 197: 1302-1306.
- Gawelek, M. A., Spataro, M., & Komarny, P. (2011). "Mobile perspectives: On iPads--why mobile?", in *EDUCAUSE Review*, 46, 2: 28-30.
- Gimenez-Lopez, J. L., Magal-Royo, T., García-Laborda, J., & Garde-Calvo, F. (2009). "Methods of adapting digital content for the learning process via mobile devices", in *Procedia Social and Behavioral Sciences*, 1, 1: 2673-2677.
- Hsu, Y., Ching, Y., & Snelson, C. (2014). "Research priorities in mobile learning: An international Delphi study", in *Canadian Journal of Learning and Technology*, 40, 2: 22.
- Ley orgánica para la mejora de la calidad educativa (LOMCE) (Ley Orgánica 8/2013, 9 de diciembre). Boletín Oficial del Estado, nº 295, 2013, 10 diciembre.
- Marks, A. M., & Cronje, J. C. (2008). "Randomised items in computer-based tests: Russian roulette in assessment?", in *Educational Technology & Society*, 11, 4: 41-50, available from <http://search.proquest.com/docview/61910453?accountid=14609>, accessed 15 May, 2015.
- McNulty, J., Chandrasekhar, A., Hoyt, A., Gruener, G., Espiritu, B., & Price, R. (2011). "Computer-based testing in the medical curriculum: A decade of experiences at one school", in *Journal of Educational Computing Research*, 45(3), 287-295.
- Nielsen, J. (1993). *Usability Engineering*. Boston: Academic Press.
- Perlman, G. (1998). *Web-Based User Interface Evaluation with Questionnaires*. Available from <http://garyperlman.com/quest/>, accessed 11 November, 2008.
- Rice, K. (2009). "Priorities in K-12 distance education: A Delphi study examining multiple perspectives on policy, practice, and research", in *Educational Technology & Society*, 12, 3: 163-177.
- Rourke, J., & Hartzman, M. (2008). "A change of direction", in *Principal Leadership*, 8, 10: 8-11.
- Saade, R. G., & Kira, D. (2007). "Mediating the impact of technology usage on perceived ease of use by anxiety", *Computers & Education*, 49, 4: 1189-1204, available from <http://search.proquest.com/docview/62056669?accountid=14609>, accessed 15 May, 2015.
- Sapriati, A., & Zuhairi, A. (2010). "Using computer-based testing as alternative assessment method of student learning in distance education", in *Turkish Online Journal of Distance Education*, 11, 2: 161-169.
- Stockdill, S.H. & Morehouse, D.L. (1992). "Critical factors in successful adoption of technology: A checklist of TDC findings", in *Educational Technology*, 32, 1: 57-58.
- Williams, D. L., Boone, R., & Kingsley, K. V. (2004). "Teacher beliefs about educational software: A Delphi study" in *Journal of Research on Technology in Education*, 36, 3: 213-229.
- Wise, S. L., Pastor, D. A., & Kong, X. J. (2009). "Correlates of rapid-guessing behavior in low-stakes testing: Implications for test development and measurement practice", in *Applied Measurement in Education*, 22, 2: 185-205, available from <http://search.proquest.com/docview/61885123?accountid=14609>, accessed 1 June, 2015.
- Wood, M. (2011). "'Closer to the ground': Pupil 'voice' and the development of knowledge about schools", in *Turkish Online Journal of Qualitative Inquiry*, 2, 1: 1-15.