British Journal of Educational Technology *Vol 48 No 2 2017* 264–278
doi:10.1111/bjet.12387

Educational games as stand-alone learning tools and their motivational effect on L2 vocabulary acquisition and perceived learning gains

José Ramón Calvo-Ferrer

José Ramón Calvo-Ferrer is a Lecturer at the Department of English Studies of the Universidad de Alicante, where he teaches different modules on English as a Foreign Language. His research interests lie in video game design for second language acquisition and testing. Address for correspondence: Dr José Ramón Calvo-Ferrer, Departamento de Filología Inglesa, Universidad de Alicante, San Vicente del Raspeig, Alicante 03690, Spain. Email: jr.calvo@ua.es

Abstract

According to different authors, computer games not only teach contents and skills, but also do so in a more efficient manner, allowing long-lasting learning. However, there is still little consensus on this matter as different studies put their educational benefits into question, especially when used without instructional support. An empirical study was conducted to measure the effect of the educational game *The Conference Interpreter* on L2 vocabulary acquisition and perceived learning gains, as compared with a non-gaming tool which replicated its contents. The results of pre-, post- and delayed tests showed that students that had access to the contents via the video game performed statistically better in the short run, found the materials more appealing and believed their vocabulary skills had developed further than those in the control group. However, the regression model showed that the actual enjoyment of the game seemed to have no effect on the students' learning outcomes, neither according to their own estimation nor as determined by testing. Of greater importance seemed to be extrinsic motivation, ie, their desire to play based upon expected learning gains, prior knowledge of tested L2 vocabulary and perceived difficulty of the educational contents.

Introduction

As different reports indicate (Entertainment Software Association, 2014; Interactive Software Federation of Europe, 2014), there has been an increase in the number of video game players in recent years. This trend seems to have paved the way for the use of video games in the classroom, leading to the design and implementation of educational computer games, also called serious games, on the part of educators, trainers and researchers. Although this may be seen as "a mad rush to pour educational content into games in an ad hoc manner" (Gunter, Kenny & Vick, 2008), different authors suggest that computer games offer a complete immersion in any type of activity, which makes them optimal tools for education (Cairns, Cox & Nordin, 2014; Warburton, 2009). Furthermore, it is argued that video games not only teach contents and skills, but also do so in a more efficient manner, allowing long-lasting learning (Nesselhauf & Tschichold, 2002; Rama, Black, van Es & Warschauer, 2012). According to Wrzesien and Alcañiz Raya (2010), educational games are powerful and effective tools since they "create personal motivation and satisfaction [. . .] reinforce skill mastery and accommodate multiple learning styles and abilities" and "provide a context for interactive decision-making." Motivation is also identified as a powerful factor in education (Anderman & Dawson, 2011), which has a positive effect especially in the field of digital game-based learning (Graesser, Chipman, Leeming, Biedenbach & Graesser, 2009).

Practitioner Notes

What is already known about this topic

- Computer games favour vocabulary acquisition in second language learning.
- Learners typically report more interest in games than in conventional instruction.
- Different studies suggest there is no difference between educational computer games and conventional instruction.
- Different authors claim that games can be detrimental to learning, especially if they do not include instructional support.

What this paper adds

- Educational games as stand-alone learning tools facilitate short-term L2 vocabulary acquisition over traditional instruction, although their effect dilutes with time.
- Differences in L2 vocabulary acquisition between educational games and conventional instruction are small.
- Cognitive engagement plays a more important part in L2 vocabulary acquisition than the feeling of fun during game play.

Implications for practice and/or policy

- Educators and game developers should consider integrating learning objectives into game dynamics for educational games to be effective stand-alone learning tools in terms of cognitive gains.
- Educational video games should include elements like progress lines or debriefing screens to raise students' cognitive engagement and maximise their impact in terms of learning gains.
- Return on Educational Investment may be maximised by providing educators with intuitive tools to create and/or adapt contents for existing games to meet specific learning needs.

Schwienhorst (2002) and Anderson, Reynolds, Yeh and Huang (2008) suggest that video games provide a realistic sociocultural context that favours second language learning. Specifically, different authors state that games are effective tools in the field of L2 vocabulary acquisition (Abrams & Walsh, 2014; Amoia, Gardent & Perez-Beltrachini, 2011; Chiu, 2013; Connolly, Stansfield & Hainey, 2011; Huyen & Nga, 2003; Pasfield-Neofitou, 2014).

However, some authors suggest that claims concerning the educational effectiveness of games are merely based upon "positive outcomes" in relation to motivation "rather than their effectiveness as standalone knowledge acquisition mechanisms" (Gunter *et al*, 2008). Brom, Preuss and Klement (2011) state that studies comparing the instructional effectiveness of games to more conventional forms of instruction show mixed results: "a) in most studies investigating motivation, learners reported more interest in simulation/game activities than in the conventional instruction but that this is not necessarily linked with better learning; b) in studies investigating cognitive performance immediately or shortly after the treatment, games were usually at least as effective as other kinds of instruction but only rarely better [. . .]; and c) that games can be detrimental to learning if they do not include instructional support." O'Neil, Wainess and Baker (2005) affirm in fact that "games themselves are not sufficient for learning," and other authors advocate for instructional support in the use of video games in order to achieve the desired learning outcome (Kirschner, Sweller & Clark, 2006; Leemkuil, de Jong, de Hoog & Christoph, 2003; Tobias & Fletcher, 2007). In line with this, Garris, Ahlers and Driskell (2002) highlight the

importance of the debriefing process in game-based education since it is "critical to making the link between game activities that may take place in a fantasy world and the application of game experiences to the real world." Other authors suggest that instructional games themselves may not motivate learners, and therefore, other sources of motivation may also be necessary to achieve satisfactory learning results (Costabile, De Angeli, Roselli, Lanzilotti & Plantamura, 2003). As regards learning outcomes perceived during game play, McMullen (1987) claims that video games have no effects on the learning outcome, although players believe they do. In line with this, Iten and Petko (2014) state that although enjoyment of a learning game has an effect on "gains in motivation to continue engaging with the subject matter being taught," it has "no discernible effect [. . .] on self-reported or tested learning gains."

Research questions and hypotheses

Based on the literature review, the following research questions were examined:

- Can video games teach educational contents in the field of L2 vocabulary acquisition without the aid of instructional support?
- Are educational video games more effective than conventional education in terms of L2 vocabulary acquisition as stand-alone learning tools?
- Do educational video games favour long-lasting learning of L2 vocabulary?
- Is intrinsic motivation in digital game-based learning a predictor of L2 vocabulary learning success?
- Does intrinsic motivation have an effect on self-reported learning gains?

To address these questions, the hypotheses of the study were formulated as follows:

H1: Students from Group A would increase their L2 vocabulary via *The Conference Interpreter* game without the aid of instructional support.

H2: Students from Group A would exhibit greater knowledge of L2 vocabulary immediately after the treatment.

H3: Students from Group A would exhibit greater knowledge of L2 vocabulary six weeks after the treatment.

H4: The more fun experienced during game play, the greater achievement in terms of L2 vocabulary acquisition students from Group A would exhibit.

H5: The more fun experienced during game play, the higher degree of perceived learning gains students from Group A would report.

Methods

Sample

The participants for this study were students of English as a Second Language in the second year of the degree in Translation and Interpreting at the University of Alicante holding a B2-C1 level of English (Common European Framework of Reference for Languages). The 65 students enrolled in this module were randomly assigned to Group A or Group B before the beginning of the experiment, although two of them did not show up and another four missed some of the sessions. In total, 59 students of Spanish nationality aged 19 and 20 years participated in the experiment (51 girls and 8 boys).

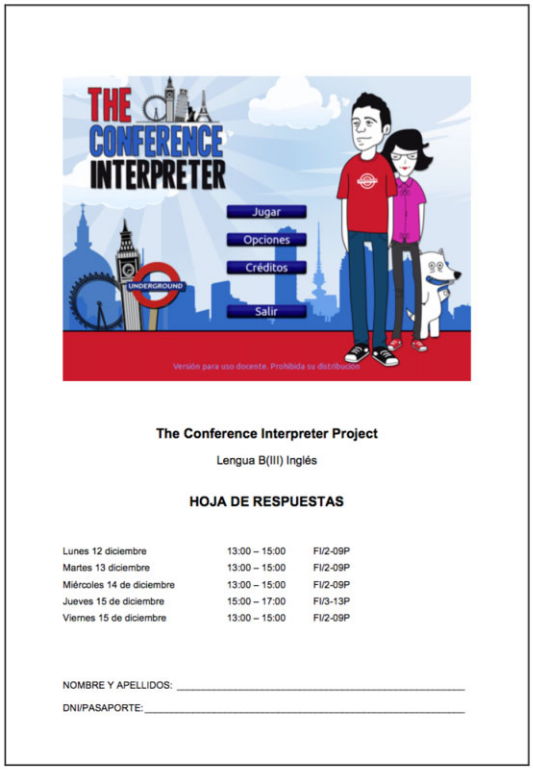
Materials

A number of studies which compare different methodologies and contents, such as explicit tuition versus computer games (see Miller & Hegelheimer, 2006; Rossiu & Papadakis, 2008; Yip & Kwan, 2006) or under different conditions (see Wijers, Jonker & Kerstens, 2008), show mixed results in terms of the effect of games in education. The purpose of our research was to isolate the gaming elements by equalling the rest of variables (stand-alone materials, identical educational contents, time of exposure to materials, etc) so that any difference in the learning outcomes could be attributed to this rather than to different approaches being applied. Therefore, although peda-

Figure 1: The Conference Interpreter *game*

Figure 2: A screenshot from the game

gogical intervention is the norm in L2 vocabulary classes, we decided to present both groups with equivalent stand-alone materials: (1) an educational game called *The Conference Interpreter* (Calvo-Ferrer *et al.*, 2013) and (2) a booklet replicating its contents. The contents were created by the researchers from online videos containing a high frequency of specialised vocabulary, namely videos of presentations of smartphones and tablets and their operating systems. *The Conference Interpreter* (Figure 1) simulates a conference that players have to simultaneously translate using the correct terminology. This is done by listening to different speakers in English and completing the translation into Spanish that appears at the bottom of the screen. Students have to select the right terms, which are elided in the Spanish text, like in conventional multiple-choice exercises (Figure 2). The game was created with the objective of developing students' terminology and interpreting skills such as divided attention, short-term memory and anticipation (see Mitchell, 2001; Pérez-Luzardo Díaz, 2005). However, as Gunter *et al* (2008) suggest, "the act of placing educational content inside a game does not guarantee that it will succeed in achieving a fun,



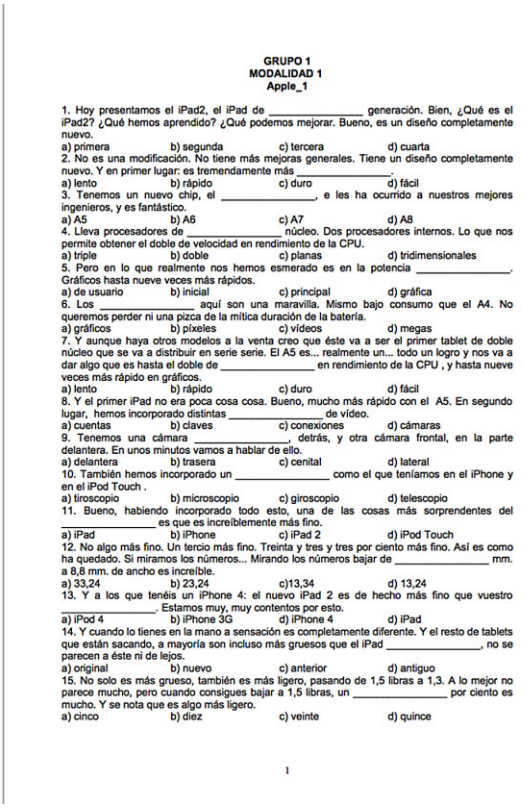


Figure 3: Booklet replicating the contents of the game

motivating experience; meeting educational goals; or being a commercial success." In line with this, the application was designed with the clear objective of being both an engaging and entertaining tool, in accordance with the principles of game design and player engagement theories (see Fullerton, 2008; Koster, 2004; Salen & Zimmerman, 2003; Schell, 2008), and a useful tool for educators, for which the RETAIN model (Gunter *et al*, 2008), the ARCS instructional design model (Keller, 1987) and the EGameFlow scale (Fu, Su & Yu, 2009) were particularly observed during its development. As different authors suggest, "to achieve an optimal learning effect, the fun of the game needs to be closely linked to the learning process" (Iten & Petko, 2014). Seventeen different versions of the video game were released and tested by second language learners other than the participants in the experiment whose feedback was analysed in order to ensure the highest possible level of user playability and engagement. The booklet, which was given to students from the control group (Group B), replicated the contents of the video game, also in the form of multiple-choice exercises that students were required to complete as they listened to the different audio files (Figure 3). Thus, while the game was developed with an instructional focus and its contents related to the curriculum, it included—as opposed to the booklet—different rewards, score tables and power-ups, provided instant feedback and its audio files were broken into different stages that students had to complete in order to move on to the next level.

Instruments

In line with the hypotheses of the study, different questionnaires were developed by the researchers to collect quantitative data regarding L2 vocabulary acquisition, student motivation and

perceived learning outcomes: (1) a knowledge test on mobile operating systems vocabulary given as pre-test, post-test and delayed test, (2) a post-test questionnaire on student motivation as developed by Guay, Vallerand and Blanchard (2000), and (3) a post-test questionnaire on students' perceived learning outcomes. The knowledge test aimed to assess students' knowledge of mobile operating systems terminology. In particular, students were asked to translate from English into Spanish 63 terms (Cronbach's $\alpha = .84$), which were directly taken from the materials provided via both the video game and the booklet. To ensure the correct understanding of the terms, sufficient context was provided, although only the translation of the actual terms was assessed. The same test was also given as a post-test and a delayed test. The post-test questionnaires on motivation and on perceived learning outcomes aimed to measure on a 5-point Likert scale (*strongly disagree*, *disagree*, *neither agree nor disagree*, *agree*, *strongly agree*) the extent to which students (1) found the materials motivating and (2) believed to have developed their L2 skills during the treatment. An exploratory factor analysis using principal components analysis with varimax rotation was conducted, and each construct was checked for reliability with Cronbach's α coefficients. The questions covered aspects related to both intrinsic motivation (five items, Cronbach's $\alpha = .95$), extrinsic motivation (five items, Cronbach's $\alpha = .81$) and perceived learning gains (seven items, Cronbach's $\alpha = .73$). As control variables, students were asked to complete a placement test (Cambridge University Press, n.d.) to measure their proficiency in English ($M = 6.56$, $SD = 1.02$) and to measure on a 5-point Likert scale whether they had found the contents of the video game easy to learn ($M = 3.93$, $SD = 1.07$) (Table 1).

Procedure

The experiment was carried out throughout 5 consecutive days. On the first day, students were given 1 hour to individually complete the pre-test questionnaire on mobile operating systems vocabulary and were randomly assigned to the different groups. Subsequently, students interacted individually with either the game or the booklet on 3 consecutive days in 2-hour sessions, under the supervision of their teachers, who, however, did not provide any help other than setting up the materials (Figure 4). On the fifth day, students were given 2 hours to complete the post-test questionnaire on mobile operating systems vocabulary, the post-test questionnaire on student motivation and the post-test questionnaire on perceived learning outcomes. No literature was found on the most adequate time to deliver a delayed test to measure long-lasting learning of L2 vocabulary. In fact, some authors measure long-term vocabulary retention several days (Bloom & Shuell, 2014), 2 weeks (Wu, 2014), 3 weeks (Groot, 2000), a month (Ellis, Tanaka & Yamazaki, 1994) and even 3 months (Elley, 1989) after the period of practice. Finally, 6 weeks after the treatment, students were given 1 hour to individually complete the delayed test in line with Brett, Rothlein and Hurley (1996), and in coincidence with their English course final exam to ensure all of the students would sit the test. Also, students were not informed about this to prevent them from revising for the delayed test nor did they have access to the learning materials other than during the period of practice with the gaming application and the booklet. All tests were completed under the supervision of a classroom teacher.

Data analysis

One-way between-subjects variance analyses (ANOVA) were carried out to investigate the effect of the materials on L2 vocabulary acquisition, motivation, and perceived learning outcomes. For the knowledge tests on mobile operating systems vocabulary (pre-test, post-test and delayed test), the total number of correct answers was calculated and presented as a total score ranging from 0 to 63 points. For the post-test questionnaires on student motivation and on students' perceived learning outcomes, the mean score of the different questions (on the Likert scale ranged 1–5) was calculated. In order to evaluate whether, along with other independent variables, motivation could account for any between-subjects differences in L2 vocabulary acquisition and perceived

Table 1: Descriptive statistics for questionnaire items

Index variable	*Sample item*	M	SD	n
Perceived learning gains (seven items, Cronbach's α = .73)	I feel that I have developed my L2 listening skills using the materials	4.32	.706	59
	I feel that I have developed my L2 grammar skills using the materials	2.53	.897	59
	I feel that I have developed my L2 vocabulary skills using the materials	4.22	.744	59
	I feel that I have learnt about the contents of the materials	3.81	.798	59
	I feel that I can reproduce the information contained in the materials	3.22	.696	59
	I feel that I can translate vocabulary from the materials into Spanish	3.36	.689	59
	I feel that I can translate vocabulary from the materials into English	2.93	.666	59
Intrinsic motivation (five items, Cronbach's α = .95)	I find the materials fun	3.71	1.068	59
	I felt motivated while I was using the materials	3.32	1.345	59
	I enjoyed using the materials	3.29	1.145	59
	I would like to use the materials again	3.54	1.222	59
	I would like to use the materials in my classes	3.03	1.273	59
Extrinsic motivation (five items, Cronbach's α = .81)	I would like to use the materials again to develop my L2 listening skills	4.25	.883	59
	I would like to use the materials again to develop my L2 grammar skills	2.49	.954	59
	I would like to use the materials again to develop my L2 translator skills	3.76	.773	59
	I would like to use the materials again to develop my L2 vocabulary skills	4.05	1.007	59
	I would like to use the materials again to obtain higher grades in my classes	3.10	1.335	59
Perceived difficulty of the materials	I find the materials easy to learn	3.93	1.07	59

Note: 5-step Likert scale: 1 = *strongly disagree*; 5 = *strongly agree*.

Figure 4: Students interacting with the game

Table 2: Descriptive statistics for pre-, post- and delayed knowledge tests by intervention group

Moment	*Group*	M	SD	n
Pre-test	Experimental	17.88	4.777	32
	Control	15.81	5.008	27
	Total	16.93	4.951	59
Post-test	Experimental	35.56	6.490	32
	Control	31.63	7.045	27
	Total	33.76	6.976	59
Delayed test	Experimental	30.41	7.224	32
	Control	27.59	8.063	27
	Total	29.12	7.684	59

learning gains, multivariate linear regressions with simultaneous inclusion of all independent were carried out. Such analyses were performed using the SPSS 22.0 statistical software with the significance level set at .05.

Results

Learning gains without instructional support

The analysis of the scores of the different knowledge tests on specialised vocabulary by students from Group A showed that there were statistically significant differences ($F[2, 62] = 172.62$, $p < .001$) between the different moments when the test was delivered, and that their effect size was high (partial $\eta^2 = .848$) according to Cohen (1988). Specifically, there were statistically significant differences between (1) pre-test and post-test ($MD = -17.688$, $p < .001$, 95% CI [−20.194, −15.181]), (2) pre-test and delayed test ($MD = -12.531$, $p < .001$, 95% CI [−15.322, −9.741]) and (3) post-test and delayed test ($MD = 5.156$, $p < .001$, 95% CI [3.070, 7.243]), which indicates that *The Conference Interpreter* game helped students gain L2 vocabulary without the aid of instructional support.

Comparison of the two groups with regard to learning effectiveness

The analysis of the scores of the different knowledge tests on specialised vocabulary showed that: (1) there were no statistically significant differences ($F[1, 57] = 2.606$, $p = .112$) between the experimental group ($M = 17.88$, $SD = 4.777$) and the control group ($M = 15.81$, $SD = 5.008$) as to knowledge of mobile operating systems vocabulary prior to the beginning of the experiment, which indicates that both groups were equivalent at that point, (2) statistically significant differences ($F[1, 57] = 4.973$, $p = .030$) resulting from the analysis of the scores of the post-test on knowledge of mobile operating systems vocabulary were found between the experimental group ($M = 35.56$, $SD = 6.490$) and the control group ($M = 31.63$, $SD = 7.045$), which indicates that students from Group A exhibited a greater level of achievement in terms of L2 vocabulary acquisition than those from Group B at the conclusion of the treatment, even though their effect size (partial $\eta^2 = .080$) was small according to Cohen (1988), and (3) no statistically significant differences ($F[1, 57] = 1.997$, $p = .163$) resulting from the analysis of the scores of the post-test on knowledge of mobile operating systems vocabulary were found between the experimental group ($M = 30.41$, $SD = 7.224$) and the control group ($M = 27.59$, $SD = 8.063$), which indicates that students from Group A exhibited a similar knowledge of mobile operating systems vocabulary to those from Group B 6 weeks after the treatment (Table 2).

Influence of intrinsic motivation on L2 vocabulary acquisition

In order to ascertain whether intrinsic motivation in digital game-based learning, ie, enjoyment of the game, could be a predictor of L2 vocabulary learning success, a linear regression was

Table 3: Influence of measured independent variables on L2 vocabulary acquisition (63 items, α = .84)

Factor	B	SE	β	T	p
Constant	4.471	10.839		.412	
Prior knowledge of tested L2 vocabulary (63 items, α = .84)	.521	.194	.384	2.686	.013
Intrinsic motivation (five items, Cronbach's α = .95)	.562	1.325	.065	.424	.675
Extrinsic motivation (five items, Cronbach's α = .81)	6.858	2.043	.537	3.356	.003
Perceived learning gains (seven items, Cronbach's α = .73)	7.740	2.642	.449	2.930	.007
Proficiency in English (*M* = 6.56, *SD* = 1.02)	1.766	.858	.279	2.058	.050
Perceived difficulty of the materials (*M* = 3.93, *SD* = 1.07)	1.591	.87	.268	1.829	.079

Note: $n = 32$; $R^2 = .590$; Adjusted $R^2 = .491$; $F(6,25) = 5.992$, $p = .001$.

Table 4: Influence of measured independent variables on perceived L2 vocabulary learning gains (seven items, α = .73)

Factor	B	SE	β	T	p
Constant	1.807	.611		2.956	
Prior knowledge of tested L2 vocabulary (sixty-three items, α = .84)	−.003	.014	−.041	−.224	.824
Intrinsic motivation (five items, Cronbach's α = .95)	.030	.087	.061	.352	.728
Extrinsic motivation (five items, Cronbach's α = .81)	.414	.138	.560	3.009	.006
Post-test on L2 vocabulary (sixty-three items, Cronbach's α = .84)	.033	.011	.569	2.930	.007
Proficiency in English (*M* = 6.56, *SD* = 1.02)	−.059	.059	−.162	−.999	.327
Perceived difficulty of the materials (*M* = 3.93, *SD* = 1.07)	−.142	.053	−.412	−2.657	.014

Note: $n = 32$; $R^2 = .481$; Adjusted $R^2 = .356$; $F(6,25) = 3.854$, $p = .007$.

calculated. As the dependent variable, Group A post-test scores were used. Overall, this model explained 59% of the variance ($F[6, 25] = 5.992$, $p = .001$), although intrinsic motivation seemed to have no significant effect ($\beta = .065$) on the learning outcomes. In contrast, extrinsic motivation, ie, desire to play based upon expected learning gains, had a positive influence on learning gains ($\beta = .537$). Students' perceived learning gains ($\beta = .449$), previous knowledge of tested L2 vocabulary ($\beta = .384$) and proficiency in English ($\beta = .279$) were also significantly influencing factors in the regression model used (Table 3).

Influence of intrinsic motivation on perceived L2 vocabulary acquisition

To analyse whether intrinsic motivation had any effect on self-assessed L2 vocabulary learning gains, a linear regression was calculated, for which the results of the questionnaire on students' perceived learning outcomes were used as the dependent variable. The regression model, which explained 48.1% of the variance ($F[6, 25] = 3.854$, $p = .007$), showed that intrinsic motivation had no significant effect ($\beta = .061$) on the dependent variable. In contrast, extrinsic motivation ($\beta = .560$), students' performance in L2 vocabulary post-test ($\beta = .569$) and perceived difficulty of the materials ($\beta = -.412$) seemed to play a part in students' self-assessed learning gains. Moreover, proficiency in English ($\beta = -.162$) and previous knowledge of tested L2 vocabulary ($\beta = -.041$) seemed to have no influence in the regression model used (Table 4).

Discussion

This paper analysed whether the educational game *The Conference Interpreter* can teach L2 specialised vocabulary without the aid of instructional support, and explored whether it is more effective than conventional education for such purposes. It also analysed whether motivation, along with other variables, is a predictor of L2 vocabulary acquisition and self-assessed learning gains.

The first hypothesis, which anticipated that students from Group A would increase their L2 vocabulary via *The Conference Interpreter* game, arose from claims which suggest that "games themselves are not sufficient for learning" unless accompanied by instructional support (see Kirschner *et al*, 2006; Leemkuil *et al*, 2003; O'Neil *et al*, 2005; Tobias & Fletcher, 2007). The results showed statistically significant differences in the pre-, post- and delayed knowledge tests delivered to students from Group A, which indicate that the educational game *The Conference Interpreter* improved their L2 vocabulary skills without the aid of instructional support, thus supporting H1. This is consistent with other findings on digital game-based L2 vocabulary learning (see Abrams & Walsh, 2014; Amoia *et al*, 2011; Chiu, 2013; Connolly *et al*, 2011; Huyen & Nga, 2003; Pasfield-Neofitou, 2014), and suggests games may be suitable stand-alone learning tools if the contents or skills being tested are adequately implemented into the game dynamics, ie, if the training objectives correspond to the skills put into practice during game play, as Belanich, Sibley and Orvis (2004) suggest.

Unlike other studies which show no difference between educational computer games and conventional instruction (see Annetta, Minogue, Holmes & Cheng, 2009; Rosas *et al*, 2003; Wrzesien & Alcañiz Raya, 2010), the post-test revealed statistically significant differences in vocabulary acquisition between the students using the game and those using the booklet. The second hypothesis, which suggested that students from Group A would exhibit significantly greater achievement in terms of L2 vocabulary acquisition than those from Group B immediately after the treatment, was thus supported. However, although statistically significant, the differences in L2 vocabulary acquisition between groups were smaller than expected. In fact, although scoring approximately twice as much as in the pre-test, students' average post- and delayed knowledge tests results were low, getting only around half of the questions right. The specialised vocabulary contained in the materials (mobile operating systems terminology), the students' background (translation and interpreting) and the limited period of practice with the materials (6 hours) are believed to have had a detrimental effect on the number of questions answered correctly and that a change in any of these variables would have resulted in higher average scores. What is more, although not significant, the differences in the pre-test might have contributed, among other factors, to the existence of such statistically significant differences in the post-test. On the other hand, such group differences in L2 vocabulary acquisition diluted in the delayed test, thus not supporting claims that video games allow for long-lasting learning more so than conventional instruction (see Nesselhauf & Tschichold, 2002; Rama *et al*, 2012). The third hypothesis, which anticipated that students from Group A would exhibit greater knowledge of L2 vocabulary six weeks after the treatment, had to be discarded. In this study, digital game-based learning seemed to favour L2 vocabulary acquisition over traditional instruction in the short run, although not in the medium–long run. This may be due to a number of possible reasons. One possible explanation is that, considering that both groups had access to identical educational contents, more training might have been necessary for the differences to become more evident. Further research is needed to ascertain whether a combination of digital game-based learning and traditional instruction (eg, by means of debriefing sessions) may improve L2 vocabulary retention in the long run as well as to investigate the correlation between the amount learned and the passage of time in order to understand for how long educational games may favour L2 vocabulary acquisition over traditional instruction and, most importantly, to analyse whether digital game-based learning may actually accelerate forgetting rates, as the results may also suggest. On the other hand, mainly a recurrent set of the terms learnt, as determined by the post-test, was forgotten when measured at the delayed test. It should be analysed whether word type may have a greater effect on the length of the recall period as compared with other variables like frequency (see Alcaraz-Mármol, 2010; Brown, 1993; Nagy, Herman & Anderson, 1985) or metacognitive awareness (see Boulware-Gooden, Carreker, Thornhill & Joshi, 2007; Cubukcu, 2008; Rasekh & Ranjbary, 2003).

The study also looked into the motivational variables that may have had an effect on tested L2 vocabulary acquisition and on students' perceived learning gains. The fourth hypothesis, which anticipated that Group A L2 vocabulary gains could be explained, among other factors, on the basis of the perception of fun playing the educational game, had to be discarded, since the analysis showed no correlation between intrinsic motivation and tested vocabulary acquisition. In contrast, extrinsic motivation and students' perceived learning gains had a greater influence than expected. These results can be interpreted to mean that students' desire to play based upon expected learning gains and anticipation of such gains plays a more relevant part in the learning that goes on in digital game-based education. This is consistent with other studies which question the actual interrelation between fun and learning (see Okan, 2003; Brom *et al*, 2011) and which suggest that extraneous cognitive load interferes with learning (Paas, Renkl & Sweller, 2003). In conclusion, although it is widely accepted that digital game-based learning favours intrinsic motivation over traditional education, the results of this study seem to highlight the need for further empirical evidence to ascertain whether intrinsic motivation may indeed result in significant learning gains.

The fifth hypothesis, which anticipated that Group A self-assessed L2 vocabulary gains could be explained, among other factors, on the basis of the perception of fun playing the educational game, also had to be discarded. According to this study, factors influencing perceived learning gains were extrinsic motivation, prior knowledge of tested L2 vocabulary and perceived difficulty of the materials. Thus, it seems that students' anticipated gains depended greatly on the potential learning benefits they could obtain from the educational game rather than on the actual fun they had during game play. Students' metacognitive awareness and learning difficulty seemed also to play a more important part than the fun factor as regards self-assessed L2 vocabulary gains, which is consistent with previous literature (Iten & Petko, 2014).

Taken together, the results of this study seem to challenge the notion of fun during game play as a determining factor in the learning that goes on in games, although intrinsic motivation has often been depicted as the cornerstone of student engagement and, ultimately, as the reason underlying the use of games in education. There are a number of plausible explanations for these results: first, the innovative elements which are intrinsic to video games (moving elements, power-ups, sound effects, scores, animations, etc) may distract players, who may end up focusing on aspects different from the learning objectives (see Papastergiou, 2009). Second, educational video games, still seen as an innovative tool in educational settings, may make students overexcited, which may also result in a distraction from their educational contents (see Wrzesien & Alcañiz Raya, 2010). Third, students motivated to learn have been typically reported to achieve better learning outcomes than those with lower degrees of motivation (McCann & Garcia, 1999; Pintrich, 2003). Thus, students willing to play based on expected learning gains may actually achieve learning objectives more easily than those more reluctant to play educational games. This seems to suggest that educational game design should keep the right balance to achieve the desired learning outcomes: on the one hand, games should be intrinsically motivating to attract and engage students. On the other hand, they should be extrinsically motivating so that students become aware of their efficacy as learning tools in order to maintain cognitive engagement during game play.

However, several limitations must be taken into consideration when interpreting these results. First, the delayed test was delivered six weeks after the period of practice, in line with previous literature (Brett *et al*, 1996). It is supposed that an earlier delivery may have brought about different results or, alternatively, may have provided a deeper insight as to when differences in L2 vocabulary learning between the gaming application and the booklet were still statistically significant. Further research should allow for spaced delayed tests in order to accurately identify

retention patterns in L2 acquisition via educational games. Second, participants who took part in the experiment were all university students aged 19 and 20. It would be important to ascertain the validity of the study with other populations such as the working and elderly. Third, the present study assessed the acquisition of L2 specialised vocabulary, namely mobile operating systems terminology. The implications of this are twofold: on the one hand, less specialised terminology may have resulted in different learning outcomes, both in the short and in the long run. On the other hand, there may be a correlation between students' level of technological affinity and learning outcomes. This is in line with some authors who suggest that, as a result of extensive contact with technology, younger individuals have developed advanced cognitive and learning skills (McNeely, 2005; Prensky, 2001; Skiba & Barton, 2006). Although there is a large body of research that argues the opposite case (Bennett, Maton & Kervin, 2008; Bullen, Morgan & Qayyum, 2011; Helsper & Eynon, 2010), future research should look into the correlation between L2 vocabulary acquisition in digital game-based environments and students' proficiency in video games.

Concluding remarks

The results of this study, along with previous literature, raise important issues for educational game developers and educators. First, it shows that games may be effective stand-alone learning tools if the learning objectives are properly integrated into the game dynamics, although its combination with complementary classroom activities may increase their potential by raising students' metacognitive awareness. Secondly, the small differences in terms of L2 vocabulary acquisition between the educational game and the booklet, along with previous studies (see Hays, 2005), may indicate that educational games do not favour cognitive performance considerably over traditional instruction. These results may question the need for educational games in terms of Return on Educational Investment. Probably, just some other teaching materials designed for identical purposes might have drawn better results in terms of Return on Educational Investment than the video game. However, since the study aimed to ascertain whether an educational game could teach L2 vocabulary both without the aid of instructional support and more efficiently than traditional education, a gaming application was devised for the purposes of the experiment. To maximise Return on Educational Investment, some authors advocate for the development of instructor-oriented authoring tools for digital game-based learning (Torrente, Moreno-Ger, Fernández-Manjón & Sierra, 2008). In addition to this, we also suggest providing educators with intuitive tools to create new contents therefore and adapt existing ones to meet specific learning needs.

Finally, the results show that intrinsic motivation by itself seems to have a small effect on learning outcomes and that students' expected learning gains may have a bigger impact on cognitive performance than the feeling of fun. These findings highlight the importance of creating mechanisms to activate students' cognitive engagement during game play. This could be done by including in-game elements to raise students' cognitive engagement, such as progress lines or debriefing screens. The results suggest that educational games should also minimise the presence of extraneous elements which could increase students' cognitive load and distract them from achieving the desired learning outcomes. This poses a great challenge for game developers, since the minimisation of in-game elements may also result in student disengagement and unsatisfactory results.

In conclusion, our findings suggest that educational games can operate as stand-alone tools for L2 vocabulary acquisition if the learning objectives are part of the game dynamics, although their effect may only be slightly bigger than conventional education and that factors like cognitive engagement or perceived difficulty may play a more important part in the learning that goes on in digital game-based learning than the feeling of fun during game play.

Statements on open data, ethics and conflict of interest

A copy of *The Conference Interpreter* video game can be downloaded for educational purposes from hdl.handle.net/10045/39015. Furthermore, free access to the data obtained during the experiment is provided via hdl.handle.net/10045/46440. All participants' data were treated confidentially and anonymously, and stored and used in compliance with the Spanish Data Protection Act (LOPD 15/1999). All the subjects understood and agreed to their participation in the study prior to the beginning of the research, and were duly informed of both the purpose of the experiment and their right to withdraw at any moment therefrom. No incentive to participate in the research was used in order to avoid a bias in sampling. None of the participants in the research considered themselves to have any disability or to be in any vulnerable circumstance prior to the beginning of the study, and all instruments and materials used were made available to all participants at the end of the study in order to minimise the effect of the experimental design. The author reports no financial or other conflict of interest relevant to the subject of this paper.

References

Abrams, S. S. & Walsh, S. (2014). Gamified vocabulary: online resources and enriched language learning. *Journal of Adolescent & Adult Literacy, 58*, 1, 49–58.

Alcaraz-Mármol, G. (2010). The role of frequency in the foreign language classroom. *Malaysian Journal of ELT Research, 6*, 167–194.

Amoia, M., Gardent, C. & Perez-Beltrachini, L. (2011). Learning a second language with a videogame. In ICT for Language Learning, Conference Proceedings 2011. Simonelli Editore.

Anderman, E. M. & Dawson, H. (2011). Learning with motivation. In R. E. Mayer & P. A. Alexander (Eds), *Handbook of research on learning and instruction* (pp. 219–241). New York, NY: Routledge.

Anderson, T., Reynolds, B. L., Yeh, X. P. & Huang, G. Z. (2008). Video games in the English as a foreign language classroom. In Second IEEE international conference on digital games and intelligent toys based education (pp 188–192). IEEE.

Annetta, L. A., Minogue, J., Holmes, S. Y. & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education, 53*, 1, 74–85.

Belanich, J., Sibley, D. E. & Orvis, K. L. (2004). *Instructional characteristics and motivational features of a pc-based game*. Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Bennett, S., Maton, K. & Kervin, L. (2008). The digital natives debate: a critical review of the evidence. *British Journal of Educational Technology, 39*, 5, 775–786.

Bloom, K. C. & Shuell, T. J. (2014). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *The Journal of Educational Research, 74*, 4, 245–248.

Boulware-Gooden, R., Carreker, S., Thornhill, A. & Joshi, R. M. (2007). Instruction of metacognitive strategies enhances reading comprehension and vocabulary achievement of third-grade students. *The Reading Teacher, 61*, 1, 70–77.

Brett, A., Rothlein, L. & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal, 96*, 4, 415–422.

Brom, C., Preuss, M. & Klement, D. (2011). Are educational computer micro-games engaging and effective for knowledge acquisition at high-schools? A quasi-experimental study. *Computers & Education, 57*, 3, 1368–1377.

Brown, C. (1993). Factors affecting the acquisition of vocabulary: frequency and saliency of words. In T. Huckin, M. Haynes & J. Coady (Eds), *Second language reading and vocabulary learning* (pp. 263–286). Norwood, NJ: Ablex.

Bullen, M., Morgan, T. & Qayyum, A. (2011). Digital learners in higher education: generation is not the issue. *Canadian Journal of Learning and Technology / La Revue Canadienne de L'apprentissage et de La Technologie, 37*, 1. http://www.cjlt.ca/index.php/cjlt/issue/view/71.

Cairns, P., Cox, A. & Nordin, A. I. (2014). Immersion in digital games: review of gaming experience research. In M. C. Angelides & H. Agius (Eds), *Handbook of digital games* (pp. 339–361). Hoboken, NJ: John Wiley & Sons, Inc.

Calvo-Ferrer, J. R., Gallego Durán, F. J., Baños Sabater, A., Reverte Bernabeu, J., Terol Sanchís, J. A., Linares Pérez, H., Santacruz Martínez, G., Márquez Rizo, J. A., Belda Medina, J. R., Llorens Largo, F. (2013). *The Conference Interpreter* [Computer software]. Retrieved 16 January 2013, from http://hdl.handle.net/10045/39015

Cambridge University Press (n.d.). Test your English. Retrieved 8 March 2014, from http://www.cambridgeenglish.org/test-your-english/

Chiu, Y.-H. (2013). Computer-assisted second language vocabulary instruction: a meta-analysis. *British Journal of Educational Technology*, *44*, E52–E56.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
Connolly, T. M., Stansfield, M. & Hainey, T. (2011). An alternate reality game for language learning: ARGuing for multilingual motivation. *Computers & Education*, *57*, 1389–1415.
Costabile, M. F., De Angeli, A., Roselli, T., Lanzilotti, R. & Plantamura, P. (2003). Evaluating the educational impact of a tutoring hypermedia for children. *Information Technology in Childhood Education Annual*, *2003*, 1, 289–308.
Cubukcu, F. (2008). Enhancing vocabulary development and reading comprehension through metacognitive strategies. *Issues in Educational Research*, *18*, 1, 1–11.
Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*, 2, 174–187.
Ellis, R., Tanaka, Y. & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, *44*, 3, 449–491.
Entertainment Software Association (2014). Essential facts about the computer and video game industry. 2014 sales, demographic and data. Retrieved 10 June 2015, from www.theesa.com/facts/pdfs/ESA_EF_2014.pdf
Fu, F.-L., Su, R.-C. & Yu, S.-C. (2009). EGameFlow: a scale to measure learners' enjoyment of e-learning games. *Computers & Education*, *52*, 1, 101–112.
Fullerton, T. (2008). *Game design workshop: a playcentric approach to creating innovative games*. New York: CRC Press.
Garris, R., Ahlers, R. & Driskell, J. E. (2002). Games, motivation, and learning: a research and practice model. *Simulation & Gaming*, *33*, 4, 441–467.
Graesser, A. C., Chipman, P., Leeming, F., Biedenbach, S. & Graesser, A. (2009). Deep learning and emotion in serious games. In U. Ritterfeld, M. J. Cody & P. Vorderer (Eds), *Serious games: mechanisms and effects* (pp. 81–100). Mahwah, NJ: Routledge, Taylor and Francis.
Groot, P. J. M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology*, *4*, 1, 60–81.
Guay, F., Vallerand, R. J. & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: the situational motivation scale (SIMS). *Motivation and Emotion*, *24*, 3, 175–213.
Gunter, G. A., Kenny, R. F. & Vick, E. H. (2008). Taking educational games seriously: using the RETAIN model to design endogenous fantasy into standalone educational games. *Educational Technology Research and Development*, *56*, 5-6, 511–537.
Hays, R. T. (2005). *The effectiveness of instructional games: a literature review and discussion*. Orlando, FL: Naval Air Center Training Systems Division.
Helsper, E. J. & Eynon, R. (2010). Digital natives: where is the evidence? *British Educational Research Journal*, *36*, 3, 503–520.
Huyen, N. T. T. & Nga, K. T. T. (2003). Learning vocabulary through games. *Asian EFL Journal*, 5, 4, 1829–1841.
Interactive Software Federation of Europe (2014). GameTrack digest: quarter 2 2014. Retrieved 10 June 2015, from http://www.isfe.eu/sites/isfe.eu/files/attachments/gametrack_european_digest_q2-14.pdf
Iten, N. & Petko, D. (2014). Learning with serious games: is fun playing the game a predictor of learning success? *British Journal of Educational Technology*. DOI: 10.1111/bjet.12226.
Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, *10*, 3, 2–10.
Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Technology Magazine: The Magazine for Managers of Change in Education*, *41*, 2, 75–86.
Koster, R. (2004). *A theory of fun for game design*. Scottsdale, Arizona: Paraglyph Press.
Leemkuil, H., de Jong, T., de Hoog, R. & Christoph, N. (2003). KM quest: a collaborative internet-based simulation game. *Simulation & Gaming*, *34*, 1, 89–111.
McCann, E. J. & Garcia, T. (1999). Maintaining motivation and regulating emotion: measuring individual differences in academic volitional strategies. *Learning and Individual Differences*, *11*, 3, 259–279.
McMullen, D. W. (1987). *Drills vs. Games—any differences? A pilot study*. Washington, D.C.: Education Resources Information Center.
McNeely, B. (2005). Using technology as a learning tool, not just a cool new thing. In D. Oblinger & J. Oblinger (Eds), *Educating the net generation* (pp. 4.1–4.10). Boulder, CO: EDUCAUSE.
Miller, M. & Hegelheimer, V. (2006). The SIMS meet ESL: incorporating authentic computer simulation games into the language classroom. *Interactive Technology and Smart Education*, *3*, 4, 311–328.

Mitchell, F. (2001). Training liaison interpreters in the classroom. In S. Cunico (Ed.), *Training Translators and Interpreters in the New Milennium* (pp. 1–8). University of Portsmouth: School of Languages and Area Studies.
Nagy, W. E., Herman, P. A. & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 2, 233–253.
Nesselhauf, N. & Tschichold, C. (2002). Collocations in CALL: an investigation of vocabulary-building software for EFL. *Computer Assisted Language Learning, 15*, 3, 251–279.
Okan, Z. (2003). Edutainment: is learning at risk? *British Journal of Educational Technology, 34*, 3, 255–264.
O'Neil, H. F., Wainess, R. & Baker, E. L. (2005). Classification of learning outcomes: evidence from the computer games literature. *The Curriculum Journal, 16*, 4, 455–474.
Paas, F., Renkl, A. & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Technology Magazine: The Magazine for Managers of Change in Education, 38*, 1–4.
Papastergiou, M. (2009). Digital game-based learning in high school computer science education: impact on educational effectiveness and student motivation. *Computers & Education, 52*, 1, 1–12.
Pasfield-Neofitou, S. (2014). Language learning and socialization opportunities in game worlds: trends in first and second language research. *Language and Linguistics Compass, 8*, 271–284.
Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*, 4, 667–686.
Prensky, M. (2001). *Digital game-based learning*. New York: McGraw-Hill.
Rama, P. S., Black, R. W., van Es, E. & Warschauer, M. (2012). Affordances for second language learning in *World of Warcraft*. *ReCALL, 24*, 3, 322–338.
Rasekh, Z. E. & Ranjbary, R. (2003). Metacognitive strategy training for vocabulary learning. *The Electronic Journal for English as a Second Language, 7*, 2, 1–11.
Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P. *et al* (2003). Beyond Nintendo: design and assessment of educational video games for first and second grade students. *Computers & Education, 40*, 1, 71–94.
Rossiu, E. & Papadakis, S. (2008). Applying online multiplayer educational games based on generic shells to enhance learning of recursive algorithms: students preliminary results. In Proceedings of the 2nd European conference on games-based learning (ECGBL). Reading: Academic Conferences Limited.
Salen, K. & Zimmerman, E. (2003). *Rules of play*. Cambridge, MA: MIT Press.
Schell, J. (2008). *The art of game design: a deck of lenses*. New York: CRC Press.
Schwienhorst, K. (2002). Why virtual, why environments? Implementing virtual reality concepts in computer-assisted language learning. *Simulation & Gaming, 33*, 2, 196–209.
Skiba, D. J. & Barton, A. J. (2006). Adapting your teaching to accommodate the net generation of learners. *Online Journal of Issues in Nursing, 11*, 2. Retrieved 10 July 2014, from http://www.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAperiodicals/OJIN/TableofContents/Volume112006/No2May06/tpc30_416076.aspx.
Tobias, S. & Fletcher, J. D. (2007). What research has to say about designing computer games for learning. *Educational Technology Magazine: The Magazine for Managers of Change in Education, 47*, 5, 20–29.
Torrente, J., Moreno-Ger, P., Fernández-Manjón, B. & Sierra, J. L. (2008). Instructor-oriented authoring tools for educational videogames. In 8th International Conference on Advanced Learning Technologies (ICALT 2008). Santander, Spain. IEEE Computer Society.
Warburton, S. (2009). Second life in higher education: assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British Journal of Educational Technology, 40*, 3, 414–426.
Wijers, M., Jonker, V. & Kerstens, K. (2008). MobileMath: the phone, the game and the math. In Proceedings of the 2nd European conference on games-based learning (ECGBL).
Wrzesien, M. & Alcañiz Raya, M. (2010). Learning in serious virtual worlds: evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*, 1, 178–187.
Wu, Y. (2014). The effect of mixed-sensory mode presentation on retaining graphic features of Chinese characters. In Z. Han (Ed.), *Studies in second language acquisition of Chinese* (pp. 126–141). Bristol: Multilingual Matters.
Yip, F. W. M. & Kwan, A. C. M. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International, 43*, 3, 233–249.