IMPROVING ENGLISH VOCABULARY LEARNING THROUGH KAHOOT!: A QUASI-EXPERIMENTAL HIGH SCHOOL EXPERIENCE

by **Martín Flores Quiroz, Ricardo Gutiérrez, Franco Rocha, María Paz Valenzuela** and **Cynthia Vilches**

Universidad Central de Chile, Santa Isabel 1278, Santiago, Chile

martin.flores @ ucentral.cl

Abstract

This research investigates the effects of the use of Kahoot! to improve English vocabulary learning in an EFL context using a quasi-experimental post-test design. An experimental and a control group from two 9th grade classes participated in the study. A pre- and post-test were applied to both groups and the scores of both groups were compared to determine whether there was any variation. The interventions in each group consisted of two lessons a week for four weeks. In particular, the results of the experimental treatment indicated an improvement regarding English vocabulary knowledge using the Kahoot! app, with a significant variation and a medium effect size. In general, this suggests the necessity of implementing new strategies with the available ICTs at hand to enhance the learning of English in Chilean classrooms.

Keywords: English vocabulary; Kahoot!; ICT; Game-based learning

1. Introduction

Although the Chilean government has enacted policies to include ICT (Information and Communication Technology) in school classrooms (Decree n°369, 2015), few institutions have included them. According to Agencia de Calidad de la Educación (2017), only 0.9% of educational institutions have advanced use of technologies in the educational field.

In conjunction with the scarce use of ICTs, the learning of English at schools concerns policy makers. In 2012, a national English test was applied, which showed that all skills were insufficiently developed. Herein, the lowest score was reading comprehension. In brief, 82% of the students cannot certify their level of English, with an average of 49 points out of 100 (70 points were needed to certify the A2 level) (Agencia de Calidad de la Educación, 2013).

As a way to tackle both these deficiencies, it is necessary to improve the strategies used to teach English by using ICTs. In this way, teachers could prepare more effective lessons for students. For that purpose, the present research aimed at finding out whether the use of Kahoot!

was effective as a vocabulary learning strategy. This research works on the assumption that by improving vocabulary, students can improve reading comprehension (Beck et al., 1983). The process of word recognition is essential for understanding an idea or a text, as the misconception of one of the components could lead to a whole new meaning (McNamara, 2012). Moreover, Farjado et al. (2012) confirmed that the more mistakes in the recognition of words students make, the more errors in reading comprehension arise after testing university students (see Wright & Cervetti, 2017, for a recent systematic review on the subject). Nevertheless, testing improvements on reading comprehension are beyond the scope of this study.

The purpose of the present research was to test the use of the ICT Kahoot! with English vocabulary learning to determine whether there exists a variation in scores comparing the pre- and post-test taken in two 9th grades. The idea was to establish whether the use of Kahoot! made any difference in contrast to the control group. Accordingly, the research question was: Does the proficiency of English vocabulary in the experimental group vary because of the use of Kahoot! as an ICT in contrast to the control group?

2. Literature review

2.1. Game-based learning

Game-based learning (GBL) is "an environment where game content and game play enhance knowledge and skills acquisition" (Qian & Clark, 2016, p. 51). In this sense, any activity, whether digital or physical, can be considered in the realm of GBL as long as game-like activities are present. In the same vein, Qian and Clark further state that these activities must have "problem solving spaces and challenges that provide players/learners with a sense of achievement" (2016, p.51).

When considering GBL in digital environments, these spaces and challenges may be placed on a continuum that can range from drill and practice to higher cognitive skills, such as critical thinking. At the same time, the chosen activity may have negative and positive effects on the students (Chen et al., 2018). Thus, in order to have a successful GBL activity, a balance must be struck between the level of challenge and the abilities of the learners to avoid, on the one hand, boredom, and, on the other, anxiety, especially if the goal is to learn vocabulary.

2.2. Vocabulary learning in digital game-based learning

As noted before, several factors must be considered to have a successful GBL activity. Particularly for digital based-game learning (DGBL), Chiu (2013) conducted a meta-analysis of

the quantitative evidence found in the literature regarding vocabulary learning through DGBL, focusing on four main aspects: treatment duration, educational level of the participants, presence/absence of teacher instruction and presence/absence of game-based learning. Regarding treatment duration, when the interventions lasted less than a month, the results were better than when they were longer. This suggests that the students may have been interested in the novelty of learning through a computer but later lost interest or felt fatigued. Considering the educational level of the participants, high-school and university students benefit more from these interventions than primary school students because the effect size is larger (d = 1.032 v. d = 0.321). Regarding teacher instruction, students showed larger improvement when they played the games without teacher assistance. Finally, this meta-analysis points out that L2 vocabulary learning has a larger effect when games are not used. Nevertheless, Chiu states that this particular finding cements the need to further study "the degree to which computer software helps learners to review vocabulary and help them acquire new words" (p. 55).

As recently noted by Cárdenas-Moncada et al. (2020), DGBL has not been extensively researched in the realm of learning English as foreign language (EFL), although there has been research on the use of DGBL to reduce anxiety and improve willingness to communicate with positive results (Reinders & Wattana, 2015). So far, the evidence is promising because the use of DGBL for EFL improves language skills development while having better results than mere educational software.

2.3. Kahoot! as a teaching tool

Although the use of Kahoot! as a teaching tool has been studied before, most research has focused on positive attitudes towards learning and motivation (Muñoz, 2016; Ramírez et al., 2017). Thus, these results suggest that Kahoot! is an adequate tool to reinforce the students' self-confidence and outlook when facing different tasks. Nevertheless, the few studies that have focused on learning, such as Iwamoto et al. (2017) in psychology, have found significant effects on the scores of the groups that have used Kahoot! as a learning tool. This conclusion is shared by Cárdenas-Moncada et al. (2020), where few studies have focused on improving EFL performance through Kahoot!, focusing mainly on other educational software (Kocaman & Kizilkaya-Cumaoglu, 2014; Alyaz & Genc, 2016). Theirs is one of the first to focus on this aspect in Chile, particularly with an adult group, while also measuring perceptions and attitudes towards the use of Kahoot!. Specifically, when it came to learning grammar, structure and vocabulary, the experimental group achieved statistically significantly higher scores than the control group. Nonetheless, there is a need to continue researching this topic.

3. Methodology

3.1. Research design

The design of this research was quasi-experimental. As Hernández et al. (2014) explain, there is an existing control over at least one of the independent variables to verify if there is any relation or effect in the dependent variable, to, then, verify the cause of it. Furthermore, the design was quasi-experimental since there was no random assignment of the participants to the control and experimental groups (Creswell, 2012). At this point, it was not possible to randomize both groups because the two 9th grade classes were conformed at the beginning of the school year and could not have any modification to their participants.

In addition to the quasi-experimental design, a control group and pre- and post-test design was included, which, as Hernández et al. (2014) state, removes the internal validity issues and allows the experimental treatment effect to be confirmed. Moreover, this is an explanatory research because the effect of the intervention, the use of Kahoot! to practice acquisition of English vocabulary, was measured to test its effectiveness (Hernández et al., 2014).

3.2. The aim of the study

The purpose of this research was to test the use of the Kahoot! tool with English vocabulary learning to determine whether there is a variation in scores comparing the pre- and post-test taken in two 9th grades. On that vein, the research question was: Does the proficiency of English vocabulary in the experimental group vary because of the use of Kahoot! as an ICT tool in contrast to the control group?

3.3. Participants and the context

The sampling of the present research was non-probabilistic. In this type of sampling, as Creswell (2012) defines it, the selection of the participants is based on their characteristics, availability and convenience. Since the national standardized English test is taken by 11th grade students, participants had to be in either 9th or 10th grade to improve vocabulary and scores in the future. The present research was carried out in a public school from Quinta Normal, Santiago, Chile. The number of subjects was twenty-eight out of thirty-one in the control group and twenty-nine out of thirty-nine in the experimental group. The total number of subjects was fifty-seven out of seventy students. The students that did not take part in one of the tests were not considered, as it would not be possible to contrast the results obtained. The study was carried out in two 9th grades and all the participants were female.

Considering the sample, the margin of error for generalization was 5.63, which means that the results can vary up to 6% approximately.

Due to only female and small sample, the generalization of the results was limited, which should be considered in future research.

3.4. Intervention materials

The materials used in the present research were divided into two categories: instruments of data collection and intervention materials. The instruments of data collection were identical pre- and post-tests for both groups that comprised four items each. These items tested vocabulary in different ways such as classifying words into categories, using vocabulary in context, matching words and rearranging scrambled sentences.

The intervention materials in the control group were a set of three worksheets with activities similar to the tests. This practice approach is often related to the Grammar-Translation Method, a traditional approach that focuses on translating from the L1 to the L2 and vice versa (Larsen-Freeman & Anderson, 2011). Therefore, the participants had to translate the meaning of the vocabulary they were learning and apply it to some activities such as multiple choice, word matching and arrange the words to form a sentence. In the experimental group, a set of five different kahoots were applied, similar to the questions presented in the tests. In these, the participants had to relate English words to images, where they could not use their L1.

3.5. Design and procedure

The present research was carried out over a period of four weeks. In the first class, both groups took the pre-test. The post-test was taken during the last class. The same researcher acted as teacher for both groups, teaching the same content; the main difference being the manner in which the content was practiced. The experimental group practiced vocabulary through Kahoot!, while the control group practiced through traditional methods, such as worksheets and exercises on the whiteboard. The pre- and post-tests were graded by two different researchers independently, in order to cross check results and avoid biases.

3.6. Data collection tools and procedures

The data were collected from a pre- and post-test. After applying the tests, a KR-20 test was conducted with a value of .88, which indicates that the pre- and post-test had high reliability (Zaiontz, 2020).

For the present research, the analysis used the Student's T test, which, as Hernández et al. (2014) state, allows to determine whether there is a significant variation between two variables or two groups. Additionally, an effect size test would be applied if the variation were significant to evaluate the size of it. As Thalheimer and Cook (2002) state, there is a specific scale that determines how big the effect of the treatment was, and that gives an insight into how large the effect was. It is important to report the effect size, since, as Field and Gillet (2010) claim, focusing only on reaching significant values while not reporting effect sizes can lead to misleading interpretations of experiments.

4. Findings and discussion

Table 1 illustrates the descriptive statistical measures of mode, mean and median obtained by the control group in the pre-test and post-test.

Table 1. Mode, mean and median of the control group

	Pre-test control group	**Post-test control group**
Mode	23	29
Mean	23	31
Median	27	29

Based on the means, it is possible to identify an increase of 35% from pre-test to post-test. Furthermore, the mode had a positive variation of 26% and the median had also a positive variation of 7.5%.

Table 2 illustrates the descriptive statistical measures of mode, mean and median obtained by the experimental group in the pre-test and post-test.

Table 2. Mode, mean and median of the experimental group

	Pre-test experimental group	**Post-test experimental group**
Mode	22	42
Mean	25	37
Median	24	38

As can be seen, the experimental group also achieved positive results. There was a positive variation of 48% between the means from pre-test to post-test. Additionally, there was also a positive variation of 91% in the mode and a positive variation of 58% in the median.

These results indicate that the experimental group increased their scores to a greater extent than the control group. Nevertheless, the previous information by itself has no statistical significance, so that Student's t-test was conducted.

The pre-test was taken to distinguish whether both groups, the experimental and control, were on equal terms at the beginning of the interventions. Table 3 illustrates the independent-samples t-test conducted to compare the pre-test mean scores in the control and experimental groups.

Table 3. Independent-samples t-test: control and experimental groups pre-test

Independent-samples t-test: control and experimental groups' pre-test	**Variable 1**	**Variable 2**
Mean	23.0714286	25.0689655
Observations	28	29
Standard deviation	8.105455737	9.346240782
Degrees of Freedom	55	
t Stat	-0.86074593	
P(T<=t) one-tail	0.19655742	
t Critical one-tail	1.67303397	
P(T<=t) two-tail	0.39311484	
t Critical two-tail	2.00404478	

There was no significant difference in the scores obtained by the control group (M=23, SD=8.1) and the scores obtained by the experimental group (M=25, SD= 9.3) in the pre-test, t (55) = -0.860, p=.196. These results suggest that before the experimental treatment began, both groups were in equal conditions, as there was no statistical difference between their score's means.

The post-test was taken to check whether there were any differences between both groups after the interventions with Kahoot! on the experimental group. Table 4 illustrates the independent-samples t-test conducted to compare the post-test mean scores in the control and experimental groups.

Table 4. Independent-samples t-test control and experimental groups post-test

Independent-samples t-test: control and experimental groups' post-test	**Variable 1**	**Variable 2**
Mean	31.3214286	37.2413793
Observations	28	29
Standard deviation	8.128840159	8.407714028
Degrees of Freedom	55	
t Stat	-2.70115077	
P(T<=t) one-tail	0.00458385	
t Critical one-tail	1.67303397	
P(T<=t) two-tail	0.00916771	
t Critical two-tail	2.00404478	

There was a significant difference in the scores obtained by the control group (M=31, SD=8.1) and the scores obtained by the experimental group (M=37, SD= 8.4) in the post-test, t (55) = -2.701, p=.004. Therefore, it can be said that Kahoot! is a useful tool that helps improve the vocabulary of the students.

Considering that the variation in the scores was significant, it was possible to identify the magnitude of the effect. In order to do it, an effect size or Cohen's d test was used. Fritz et al. (2012) state that the effect size allows researchers to focus on a quantitative description, generally more interpretable and comparable to other studies, because statistical significance can always be achieved with a large sample.

The intervention had a medium size effect (d = .716) (Cohen, 1988), well within Hattie's zone of desired effects (2009). This medium effect size means that almost 76% of the participants in the control group achieved a score lower than the average participant in the experimental group (Coe, 2002). Therefore, the difference in vocabulary learning achieved by the Kahoot! intervention is not only significant but also important.

5. Conclusion

As stated in the literature review, there is still a need to investigate whether the use of DGBL improves language learning, particularly, vocabulary, not only because of the conclusions reached by Chiu's meta-analysis (2013), where vocabulary learning was more effective when no games were involved, but also because most studies that have used Kahoot! as a teaching tool have focused on motivation and attitudes towards learning.

The findings of this study are consonant with what other researchers have concluded so far: Kahoot! can be used to improve learning. From a general learning perspective, these results agree with Iwamoto et al. (2017)'s conclusions that Kahoot! can be used to achieve significant differences in scores when compared to traditional learning methodologies. As presented in the results section, the higher results achieved by the experimental value are significantly different from the experimental group (t (55) = -2.701, p=.004). Particularly in relation to learning English as a foreign language, these results provide further evidence to Cárdenas-Moncada et al. (2020)'s findings that Kahoot! can be used to improve English vocabulary, not only in a group of adults as they did.

Interestingly, while most studies with Kahoot! have focused on university students, this one has focused on high school students, achieving positive results in vocabulary learning. As stated by Chen et al. (2018), the possibility to interact and receive immediate feedback are

features that allow students to avoid boredom and anxiety when it comes to DGBL. These two characteristics are present in Kahoot!, which clearly relate to the positive impact it has on motivation as described in previous studies, and also on how the participants of this study engaged and achieved higher scores. Additionally, being able to compete among peers while keeping a live feed of the scores further engages learners (Qian & Clark, 2016).

Regarding the impact of this study, English vocabulary learning scores were improved in the experimental group in a significant way by also studying a demographic that had not been considered in previous studies with Kahoot!. This opens the possibility to improve vocabulary learning and, in turn, test if this type of improvement may improve reading comprehension (Wright & Cervetti, 2017). This is a positive outcome especially if we consider that, as mentioned before, reading comprehension in English is the weakest skill in Chilean high school students. At the same time, the intervention had a medium effect size (d=.716), something that had not been reported in previous EFL studies and that allows this study to be compared with other studies in future meta-analyses.

Even though the results were positive, it is important to mention that this study did have some limitations that need to be considered in future studies. First, the school where this study took place was an all-female public high school. Secondly, the sample size was rather small (<30 per group), which could generate problems when generalizing these results to larger populations. Thirdly, this school only represents a very particular sociodemographic group, lower middle class. Gender and socioeconomic status must be considered in future research. Fourthly, this research only focused on the use of a particular type of ICT. Therefore, future research should identify the available ICT resources in the Chilean classroom and generate interventions based on them. In addition, a larger sample should be considered, while also reporting the effect size of the intervention. Finally, future research should study if, by improving vocabulary through Kahoot!, learners also improve their reading comprehension by testing this skill before and after a vocabulary-focused intervention, which was beyond the scope of this research.

Acknowledgement

We would like to thank the two anonymous reviewers for their constructive comments on my paper.

References

Agencia de Calidad de la Educación. (2013). *Informe de Resultados SIMCE inglés III medio 2012 para Docentes y Directivos.* http://archivos.agenciaeducacion.cl/biblioteca_digital_historica/resultados/2012/inserto_2012.pdf

Agencia de Calidad de la Educación. (2017). *Percepciones acerca del uso de Tecnologías de Información y Comunicación (TIC) y los aprendizajes de los alumnos de enseñanza media*. http://archivos.agenciaeducacion.cl/06_Tics_y_aprendizajes.pdf

Alyaz, Y. & Genc, Z. (2016). Digital game-based language learning in foreign language teacher education. *Turkish Online Journal of Distance Education, 17*(4), 130-146. https://eric.ed.gov/?id=EJ1116837

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*(4), 506-521. https://doi.org/10.1037/0022-0663.74.4.506

Cárdenas-Moncada, C., Veliz-Campos, M., & Veliz, L. (2020). Game-based student response systems: The impact of Kahoot in a Chilean vocational higher education EFL classroom. *CALL-EJ, 21*(1), 64-78. http://callej.org/journal/21-1/Cardenas-Veliz-Veliz2020.pdf

Chen, M., Tseng, W. & Hsiao, T. (2018). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology, 49*(1), 69-77. https://doi.org/10.1111/bjet.12526

Chiu, Y. (2013) Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology, 44*(2), 52-56. https://doi.org/10.1111/j.1467-8535.2012.01342.x

Coe, R. (2002). It's the Effect Size, Stupid. What Effect Size Is and Why It Is Important [Paper]. The British Educational Research Association Annual Conference, Exeter, U.K. http://www.leeds.ac.uk/educol/documents/00002182.htm

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.) New York: Routledge.

Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research (4th ed.).* Boston: Pearson.

Decreto n°369. Diario Oficial de la República de Chile, Santiago, Chile, December 2, 2015.

Farjado, A., Hernández, J., & González, A, (2012). Acceso léxico y comprensión lectora: un estudio con jóvenes universitarios. *Revista Electrónica de Investigación Educativa, 14*(2), 25-33. http://redie.uabc.mx/vol14no2/contenidofajardoetal.html

Field, A. & Gillet, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*(3), 665-694. https://doi.org/10.1348/000711010X502733

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2. https://doi.org/10.1037/a0024338

Hattie, J. (2009). *Visible Learning*. Abingdon: Routledge.

Hernandez, R., Fernandez, C, & Baptista, P. (2014). *Metodología de la Investigación* (6th ed.). México D. F.: McGraw-Hill/Interamericana.

Iwamoto, D., Hargis, J., Taitano, E., & Vuong, K. (2017). Analyzing the efficacy of the testing effect using Kahoot on students. *The Turkish Online Journal of Distance Education, 18*(2), 80-93. https://eric.ed.gov/?id=EJ1145220

Kocaman, O., & Kizilkaya-Cumaoglu, G. (2014). The effect of educational software (DENIS) and games on vocabulary learning strategies and achievement. *Education and Science, 39*(176), 305-316. https://dx.doi.org/10.15390/EB.2014.3704

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and Principles in language teaching* (3rd ed.). New York: Oxford University Press.

McNamara, D. S. (Ed.). (2012). *Reading Comprehension Strategies: Theories, Interventions, and Technologies*. New York: Psychology Press.

Ministerio de educación (2015) *Bases Curriculares 7º básico a 2º medio*. https://media.mineduc.cl/wp-content/uploads/sites/28/2017/07/Bases-Curriculares-7%C2%BA-b%C3%A1sico-a-2%C2%BA-medio.pdf

Muñoz, M. (2016). *Las TICs en Educación: Kahoot! como propuesta de gamificación e innovación educativa para educación secundaria en educación física* [Unpublished Master's dissertation]. Logroño: Universidad Internacional de La Rioja.

Qian, M. & Clark, K. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behaviour, 63*, 50-58. https://doi.org/10.1016/j.chb.2016.05.023

Ramírez, A., Arciniega, A., Iriarte, A. & Arriaga, M. (2017) Educational applications for teaching: Case study Kahoot!. *Revista EDUCATECONCIENCIA, 16*(17), 139-149. http://tecnocientifica.com.mx/educateconciencia/index.php/revistaeducate/article/view/391

Reinders, H. & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL, 27*(1), 38-57. https://doi.org/.1017/S0958344014000226

Thalheimer, W. & Cook, S. (2019, October 9). *How to calculate effect sizes from published research articles: a simplified methodology*. Doctoral Research Methods – Content. http://www.bwgriffin.com/gsu/courses/edur9131/content/Effect_Sizes_pdf5.pdf

Wright, T. & Cervetti, G. (2016). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly, 52*(2), 203-226. https://doi.org/10.1002/rrq.163

Zaiontz, C. (2019, July 4). *Kuder and Richardson Formula 20*. Real Statistics Using Excel. http://www.real-statistics.com/reliability/internal-consistency-reliability/kuder-richardson-formula-20/