Exploring High Technology Gamification in the Spanish CLIL EFL Classroom: A Case Study

Stephanie Tenepere Cuaderno

Máster en Lingüística Aplicada al Inglés

MÁSTERES
DE LA UAM
2018 - 2019

Facultad de Filosofía y Letras

UAM Universidad Autónoma de Madrid

Exploring High Technology Gamification in the Spanish CLIL EFL Classroom: A Case Study

Stephanie Tenepere Cuaderno

Tutor: Margarita Vinagre

Masters Degree in English Applied Linguistics
Facultad de Filosofía y Letras
Universidad Autónoma de Madrid

Academic Year: 2018/2019

Acknowledgements

First, I would like to give a warm thanks to all of my amazing students who helped make this project possible in the first place. Working with them as their English Language Assistant during the 2018-2019 academic year made me remember how crazy, exciting, and inspiring teaching and working with kids abroad can be. I also give thanks to my fellow teaching staff members who made me feel part of the team and always showed the utmost kindness and support towards my pursuit of this masters degree.

Furthermore I would like to thank my supervisor Professor Margarita Vinagre, who first introduced me to the use of technology in the foreign language learning classroom. Her course challenged me to explore outside of my academic/teaching comfort zones, and I thank her for her encouragement and patience. I could not have written this dissertation without her professional help and guidance.

Last but not least, I would like to express my love and gratitude for all of my recklessly fun and caring family members who have been by my side through every wild decision I've made in life up until now. I thank them for always believing in me and reminding me that I am strong and capable of pursuing anything that I set my mind to.

In loving memory of Sylvester, Bebe, Lola, Austin, and Grandpa Stan. Life is not forever, but neither are goodbyes.

Stephanie Cuaderno
September 2019

Abstract

Due to the European Union's current educational policies on promoting bilingualism, Spain has invested a lot of time and energy to implement English as a Foreign Language (EFL) into its national educational curriculum, particularly through a teaching method known as Content and Language Integrated Learning (CLIL). However, Spain's progress with CLIL and EFL has been overall slow, and there is obviously room for improvement. In response, this study investigated the potential of gamified learning as a modern alternative teaching methodology fit to support CLIL EFL learning goals. The researcher explored gamified learning in two specific contexts, high tech gamification and low tech gamification, with aims to test if both kinds would prove to be effective in increasing students' EFL test scores in general, and if so, whether high tech gamification would be significantly more effective from a statistical perspective. A quantitative study with a pre/post test design was used on an experimental and a control group consisting of Spanish 6th grade CLIL EFL students learning A2 level of English vocabulary in order to compare how the students performed after being exposed to one of these forms of gamification. The main findings of this study were that both high and low tech gamification increased average test scores, although only the results of high tech gamification were statistically significant.

Keywords: gamification in education, gamified learning, Content and Language Integrated Learning (CLIL), English as a Foreign Language (EFL), teaching methodologies

Index.

"This is the prescription: if you have children or you work with children…get into the game with your kids…Don't fight the game trend. Become one with the game. Enter the game. Understand it. Understand the dynamic of how your children play the games that they play. Understand how their minds work from the context of the game outward, rather than from the world outside inward." –Gabe Zichermann, TEDxKids 2011

1. Introduction

English as a Foreign Language (EFL) has come to be a dominantly learned language subject across many schools around the world. In the European Union, for instance, it is the most commonly learnt foreign language at both the primary and secondary school levels, whether compulsory or not (Baidak, Balcon, & Motiejunaite, 2017). Despite efforts in the EU to achieve national proficiency in EFL, effectively teaching the subject has not been a straightforward path for many countries, such as Spain, which has motivated the rise of new 21st century teaching methodologies that are challenging traditional classroom settings. Such methodologies include Content and Language Integrated Learning (CLIL) and gamified learning. While CLIL has established substantial roots in Spanish EFL education, the potential role of gamified learning in still in blossom. However, both approaches are yet to be fully explored and understood by researchers. Moreover, little has been studied on how the two interact together as a combinative approach to teaching EFL.

1.1. Brief Background of CLIL

CLIL emerged in the early-mid 90s and was endorsed in 2003 by the European Commission as an adequate approach to help meet the EU's foreign language learning objectives. In the EFL context, CLIL entails the usage of English as a Medium of Instruction (EMI)[1] to teach not only EFL, but also other content classes (e.g. sciences, math, art, etc.). Data from the European Commission reports that most countries in the EU offer some form of CLIL (in which English CLIL is assumed, but not directly specified), but it is still not considered by any means a widespread teaching method (Baidak et al., 2017). Like any new teaching trend, it currently faces many challenges as it tries to gain a greater foothold in language education. It has been both praised and criticized by those in the field, but ultimately, researchers and EFL instructors are still learning how CLIL practices can be improved in the classroom and what measures can be used to maximize the language learning experience.

[1] English as a Medium of Instruction (EMI) has become widely used in Higher Education pedagogy and research while CLIL is associated to primary and secondary education

1.1.1. CLIL in Spain

Spain, in particular, is considered to be one of the countries fostering many ripe opportunities for CLIL research, which is reflected by its widespread regional initiatives to seriously adopt the methodology in EFL education, as well as in other foreign/second languages (see *CLIL in Spain,* Lasagabaster & Ruiz de Zarobe, 2010). These initiatives began taking root in 1996 when the Spanish Ministry of Education, Culture and Sport and the British Council founded the nationwide state program called the Bilingual and Bicultural Project. Thanks to this project, findings report that by 2016 approximately all of Spain's primary and secondary students were learning English (Eurostat, 2018).

Interestingly, despite Spain's overall national commitment to fostering EFL education through CLIL over the past two decades, it has traditionally held unimpressive results regarding its population's overall English language proficiency level compared to other countries in the EU. According to EF Education First's English Proficiency Index of 2018, for instance, which measures the English proficiency levels of adults in EFL learning nations, although Spain overall ranks at a moderate language proficiency level globally, it holds one of the top three lowest scoring profiles in the EU, barely pulling just fractions of a point ahead of Italy and France. This phenomenon led the researcher to acknowledge not only that there are obviously many obstacles to using CLIL in EFL learning in Spain, but also and more importantly that solutions need to be offered and investigated.

1.2. Background of Gamification in Education

Gamification as a teaching practice in general emerged approximately two decades after CLIL. In comparison, it is still in the exciting phases of development, and has a while to go before being more widely considered or recognized as a proven teaching methodology. Regardless, it is often commended for its capacity to motivate students to stay engaged in learning more than traditional instruction can. However, understanding gamification and how instructors can effectively harness it in the classroom requires more time and research.

1.2.1. What is gamification and gamified learning?

There are many existing definitions for the term ‘gamification’, but one of the most common definitions referred to throughout relevant literature is usually a slight variation of Wikipedia’s (2019) version: “the application of game-design elements and game principles in non-game contexts.” This generic definition is useful because it renders gamification as a process that is applicable to many different contexts in real life. As suggested by its name, ‘gamified learning’ is thus interested in the application of gamification in the field of education, also referred to as the gamification of learning and instruction, which according to literature reviews (e.g. Caponetto, Ott, & Earp, 2014; Hamari, Koivisto & Sarsa, 2014) is one of the most popular ways gamification has been empirically studied in general. Drawing from the elements of the first definition, gamified learning in this study can be simply described as the application of game-based elements to educational activities in order to promote learning.

1.2.2. What kinds of gamification are there?

The usage of technology is often an assumed feature of gamification in its 21st century context, but technology is not necessarily a required component to create a gamified learning experience, nor is it always readily available to everyone in today’s society, despite its existence. Gamification can be divided into different levels regarding its use of technology: low tech and high tech. Understandings of both types are relevant for this study.

The term *high technology (high tech) gamification* simply refers to any gamified learning activities that consistently rely upon the use of some form of basic modern technology (e.g. computers, Internet/Wifi, tablets, smart phones, applications, websites, online platforms, etc.) in order to be fully functional. More specific examples include online learning websites or apps such as Kahoot, Socrative, Quizlet, or Duolingo—all which require either a computer or mobile device and an Internet connection in order to function.

In contrast, *low technology (low tech) gamification* refers to any gamified learning activities that do *not* rely upon or are based on the use of any basic modern technology for full functionality. This definition falls in line with the Merriam-Webster dictionary’s 2019 English Language Learners definition of ‘low-tech’ as “not using new electronic devices and technology: technologically simple.” In correspondence, based on the first part of this definition, low tech gamified learning implies that activities are conducted primarily through the use of traditional

learning materials like those found in a typical classroom (e.g. paper, chalkboards, pencils, etc.). The second part, “technologically simple” implies that the instructor may use basic modern technology at some point of the gamification process, however in this case, the use of technology aids in the implementation of the gamified activity without actively transforming the gamified learning process. Gamification in a low-tech setting may result, for example, in the use of paper-based gamified activities. To explain, manually creating enough learning materials for a class of students often consumes much labor and time, so simple technology may be used to aid in the production process (e.g. the use of a computer and a printer to print quizzes or flashcards in bulk). Low tech gamification may be a ready alternative when modern technology is available for staff, but perhaps not widely available for an entire classroom or school of students.

1.2.3. Why is gamified learning relevant in the 21st century?

While of course finding methods to motivate students has always been a natural concern for most educators, gamification is especially alluring in the sense that it seems to complement the norms of today’s youngest generation of students, labeled by historians as “Generation Z” or simply put, “Gen Z.” From an educational perspective, Gen Z is unique in that it is the first generation of students for whom technology has been by default “always on” (e.g. regular access to TVs, smartphones, Wifi, social media, tablets, etc.) (Dimock, 2019). Consequently, gamification, which as mentioned before is often automatically connoted in its high tech form, is seen as a modern approach that has the potential to transform the EFL and CLIL classroom and motivate students to learn in a way that is both fun and relevant to their daily interests, technological affinities, and skills.

1.3. Gaps in this Body of Knowledge

The optimistic attitudes towards the motivational qualities of gamified learning as well as its conceived suitability to 21st century needs have contributed to the rapid growth of it as a research topic, but not necessarily in the most directly, scientifically meaningful ways. For instance, Dicheva, Dichev, Agre and Angelova (2015) and Dichev and Dicheva (2017) discovered approximately 6600 studies specifically on gamification and education in online databases between the years 2010-2015, with the majority of those studies appearing by 2013. However, despite this high quantity of papers produced in such a short time span, the existing

research on empirical studies on this topic was actually very limited; only 85 of these studies were identified as based on empirical data. Similar disproportionate results were found by Hamari et al. (2014) who had also published a similar review of studies around the same time frame, but on the topic of gamification in relation to motivational affordances; only 9 of these empirical studies were found in the context of education and learning. Sailer and Homner (2019) conducted the latest review update from all years up until 2017, and only found a total of 45 empirical studies that were classified as solely quantitative studies on gamified learning.

Despite the few total studies reviewed up until 2017, the research on gamification and education has been rather diverse. While diversity in any research field is usually enlightening, a small base of empirical studies that comprises of various subtopics actually complicates broader understandings of the main topic at hand, especially during the early phases of research. The empirical literature of gamification in education for instance is highly diversified by many factors: subject, education level, types of research study, types of game-elements, etc. Not many studies target the same experimental variables, which although does not invalidate the findings, it isolates them as non generalizable case studies. At the same time, there are certain biases for specific areas of gamification in education; for example: languages are a poorly represented subject (e.g. Dichev & Dicheva, 2017); university-level populations skew the population samples (e.g. Caponetto et al., 2014; Subhash & Cudney, 2018); and gamified designs often favor the implementation of the most basic game element, reward structures (e.g. Hamari et al., 2014; Majuri, Koivisto & Hamari, 2018).

In sum, the disproportionate levels of empirical studies to total studies, the uncontrolled mixing of variables, and the observed biases in gamification studies have consequently left many gaps in the literature, gaps which extend over to even narrower scopes of gamification in education, such as gamification in EFL and/or CLIL contexts. As a simple demonstration, a basic search on Google Scholar (one of the most popular and largest online academic databases) shows about 1000 hits for gamification in EFL, and merely 234 for gamification in CLIL since 2015. Based on previous proportions of actual empirical studies discovered and reviewed in gamification and education in general, findings of a significant proportion of empirical data for gamification in the narrow scopes of CLIL and EFL are not likely. The resulting impression is that while the interest in gamification in education, EFL and CLIL clearly exists, there is still an evident need not only for more substantial empirical studies, but also more consistently categorized paths of research.

1.3.1. Addressing the Gap

Addressing the many challenges of EFL learning under CLIL contexts in Spain is a complex and regional topic that will require several different solutions. Simultaneously, due to its newborn status in the research field, gamified learning in CLIL contexts still has many unchartered waters, not only in Spain, but also in other CLIL classrooms across the EU. This study was intended to investigate the proposal that gamified learning could prove to be one of the solutions to help improve EFL learning in Spanish CLIL classrooms. To narrow the scope even further, this study was aimed to address a relatively unexplored factor in gamification and education studies as a whole: the element of high technology as a significant component of effective gamified learning. Granted that this study adds to the diversity of the entire literature base, conducting an empirical study on this specific topic was also intended to be used as an opportunity to slowly help fill in a couple of the existing gaps in gamification and education literature, such as the need for more studies at the K-12 education level and for studies aimed at languages/EFL subjects.

1.4. Purpose and Scope of the Study

The main purpose of this study was to present new research findings of gamification in education. To do so, this study was aimed at investigating the comparative effectiveness of two distinct modes of gamification applied to Spanish CLIL and EFL teaching: high tech gamification and low tech gamification. In doing so, the researcher wished to explore if and to what extent high tech gamification was more effective than low tech gamification, and also, as a supporting aim, to broadly observe whether high tech gamification, as it is normally assumed to be in the modern day context, seemed to share a positive relation to students who were born into a highly technologically-oriented age, Gen Z. These goals were driven by the desire to search for alternative solutions to strengthen the Spanish CLIL EFL experience. This dissertation reports upon a case study with a quantitative research design that compared the effective use of high tech and low tech gamification in an EFL classroom to improve the test scores of 6th grade students from a Spanish CLIL-based primary school. The scope of this study was mainly selected out of convenience and the availability of technological resources, space, and participants.

1.5. Research Questions

In light of the researcher's motivations for this study, the following research questions and hypotheses were posed:

- **Research Question/Hypothesis 1:** Does gamification result in improved EFL learning test scores for 6th grade Spanish CLIL students, regardless of the level of technology implemented (high tech/low tech)? *Both high tech and low tech gamification will improve EFL learning test scores of 6th grade Spanish CLIL students.*
- **Research Question/Hypothesis 2:** Is high tech gamification more effective than low tech gamification in improving EFL test scores for 6th grade Spanish CLIL students? *High tech gamification results will be statistically significantly higher than low tech gamification results in improving the students' test scores.*

1.6. Significance of the Study

There were many practical reasons to investigate the answers to these research questions. Firstly, despite the enthusiasm and commitment that Spanish educational authorities have dedicated toward CLIL as a stepping stone to EFL education, it is apparent from the data that they are still falling short of producing a significant impact on Spanish EFL learners' performance. This calls for an overall need for improvement in CLIL in Spain, and this study is meant to contribute to the search for modern, alternative solutions, which ultimately directly benefits CLIL instructors and Spanish EFL students as a whole in Spain. Secondly, the results of this study are meant to help current and future CLIL instructors recognize the generational gaps between them and their students. In Spain specifically, as of 2016, approximately only 9% of primary and secondary teachers were under the age of 30 (OECD, 2018). This implies that most of Spain's CLIL EFL teachers come from generations that were significantly less-technologically dependent or savvy, and thus less likely to be incorporating teaching methods that could fully maximize the experience of EFL learning in the 21st century. The results of this study were intended to potentially demonstrate how other modern teaching methods, such as 21st century high tech gamification, could better aid students in achieving their EFL learning goals, and encourage teachers to renovate their methods. Thirdly, finding evidence of the effectiveness of high tech gamification could offer useful insights about the role of

technology in the learning experiences of Gen Z students, and promote CLIL instructors, who are interested in enhancing their practices through gamified learning, to consider implementing more high tech gamification whenever feasible. Lastly, the findings of this study were intended to make a new contribution to the uneven literature of empirical studies in gamification in education/Spanish CLIL contexts.

1.7. Outline of the Content and Structure of this Dissertation

The remainder of this dissertation will address the following main sections: a literature review on gamification in education, as it pertains to the delimitations/scope of the conducted study, the methodology used, an analysis and discussion of the final results, and a concluding summary that includes a review of the study's inherent limitations and suggestions for which direction similar future studies may take.

2. Literature Review

The following chapter is a literature review of gamification in education. The first aim is to discuss the basic aspects of gamified learning as a theory. The second aim is to discuss some of the main background theories and principles supporting gamified learning. The final aim is to present a summary and critique of past research studies that are relevant to the main variables presented in the research study.

2.1. Gamified Learning/Gamification in Education: Definition

Landers (2015) defines gamified learning as "the use of game attributes…outside the context of a game *with the purpose of affecting learning-related behaviours or attitudes*" (p.1, emphasis added), which is a simple extension of the generic definition of gamification discussed earlier as "the application of game-design elements and game principles in non-game contexts." It is important to acknowledge that "game-based learning" is distinct from gamified learning, or gamification in education; gamified learning is "adding game inspired elements to your course" and game-based learning is "using games to meet learning outcomes" (Isaacs, 2015). Gamified learning does not imply transforming a learning activity into a full on game, but rather is the intentional application of specific game-based elements to an activity when seen as appropriate. From a theoretical educational perspective, gamified learning is a potentially valuable new teaching method based on its abilities to optimize students' engagement in learning. It proposes that the creative use and combination of game-based elements can be adapted into an educational context in order to form a truly fun and engaging learning experience, which will consequently positively affect both students' psychological and behavioral outcomes in the classroom. It is therefore necessary to define what exactly is meant by the umbrella term 'game-based elements,' a term that is often conflated or used interchangeably with other gamification terms like 'game principles,' 'game mechanics,' or 'game variables.' The significance of all of these terms is mainly to highlight researchers' observations that there are special qualities or characteristics of games that make them distinguishable from broader similar concepts, such as play.

There are many types of games and they all use different game-based elements in ways unique to themselves. Some games may incorporate just a few elements into their design while others strategically combine a complex variety of them. As a general rule of thumb, the

more game-based elements that are incorporated into a game, the more complex it becomes. Some authors claim that this complexity fosters an ideal atmosphere for the most effective learning (e.g. Kapp, 2012). However, based on the existing research in gamification, others caution that the use of any set number of elements, nor the specific combination of elements, has been properly validated (Dichev & Dicheva, 2017). Moreover, certain game-based elements have been favored over others, such as the frequent usage of reward structures in studies observed by Dichev and Dicheva (2017), and Subhash and Cudney (2018). Nonetheless, the popularity of a given element should not be regarded as an accurate measure of effectiveness. For instance, Nicholson (2015) suggests that the use of reward structures alone is not a sustainable strategy to effectively motivate players, and argues that other game-based elements should be creatively implemented to create long-term, meaningful gamified learning experiences.

While games will always vary by design, it may be broadly assumed that all games attempt to use their elements to contribute to the larger goal of sustainably engaging the player enough to keep returning back to the game, even if the game has already been won once before. This is usually what distinguishes a successful game from a poor one. Table 1 summarizes many of the core types of game-based elements often encountered in games, all of which function differently towards engagement (adapted from Kapp, 2012, 25-50).

Table 1. Game-Based Elements

Game-Based Element	Description	Example
1. Abstractions of concepts and reality	The modeling of games after real world experiences, but with reduced complexity; allows players to grasp concepts faster by providing clear cause and effect relationships.	The Sims: a life simulation game that models how humans progress through major life stages, such as starting a family, getting a career, building a house, etc. Reduces the concept of time.
2. Goals	Adds meaning or purpose to a game and provides a measurable outcome.	Chess: a game in which players win by forcing their opponent's King into checkmate. Progress is usually measured by how many of the opponents' pieces have been eliminated off the board.

3. Rules	A set of guidelines of how to play a game; designed to limit player actions to maintain a sense of order.	Checkers: a game in which pieces can only move in diagonal directions, not freely across the grid.
4. Modes of Play **a. Conflict** **b. Competition** **c. Cooperation**	Controls how players socially interact in the game. a. Players must actively defeat each other to win. b. Players compete against each other to win, but without directly interfering in each other's progress. c. Players work together to achieve the goal.	a. Arcade 1 vs 1 combat games-players must attack the other player until defeat. b. Bowling-players independently compete to knock over bowling pins with a ball. c. Team sports-games that require multiple players and teamwork to win.
5. Timers	Induces stress/pressure on players to use their skills to achieve the goal faster; provides a measurement of progress.	Trivia games: players must correctly answer questions before the time runs out.
6. Reward Structures	Incorporates a structure that extrinsically incentivizes players through some kind of reward.	Point systems, rewards, badges, bonus levels, treasures, extra lives, leaderboards, etc.
7. Feedback	Features that provide cues of player progress; informs them when they make mistakes; gives them opportunities to experiment with how to correct their actions.	Life/energy/power level bars, monitors of number of lives/tries left, active commentary, etc.
8. Levels	Adds complexity to a game through different types of level; allows proper spacing of the game content; provides scaffolding for obtaining necessary game knowledge and playing skills.	Game levels (progressive stages of increasing difficulty), playing levels (easy/intermediate/advanced), and player levels (newbies/veterans).
9. Storytelling	Weaves a story narrative into the game via protagonists, villains, a plot, etc.; provides fantasy.	Clue: a game with a murder mystery narrative involving various game character profiles.

10. Curve of Interest	How the events of a game are sequentially organized to be increasingly more engaging; hooks player attention at the beginning and then continues to build interest until the climax/goal is reached.	Risk: a game that starts off slow with the distribution of territories and the collection of armies; interest in the game builds up as fighting takes place until the eventual conquest of the world which signifies victory.
11. Aesthetics	Visual elements that are designed to make the game environment more detailed and attractive.	Candy Crush: a mobile game app that incorporates the use of bright, sparkling candy pieces that capture the eye.
12. Replay or Do Over	A feature that gives players permission to fail and try again; fosters curiosity to explore the game, try different strategies, and achieve a greater sense of fulfillment when the goal is finally reached.	Try again/play again settings.

2.2. Underpinning Theoretical Perspectives behind Gamification in Education

Much of the justification for gamified learning is based on a mix of theories primarily about motivation, while others draw upon theories that address different aspects of learning. The following section aims to discuss these theories and how they have shaped gamification in education.

2.2.1. Motivation Theory

As explained by Kapp (2012), motivation theory is essentially about understanding what drives a person to do something, or to work towards a given goal. There are two main types of motivation: intrinsic (internal) and extrinsic (external) motivation. Intrinsic motivation is the concept that a person's motivation is driven from within; he/she chooses to do something because it appeals to an inner desire or is personally satisfying in some way, for example. Their willingness to achieve a particular goal functions independently of any expected consequences if that goal is achieved. On the other hand, extrinsic motivation is driven by external motives; a person is motivated to do something not for the sake of his/her personal interest in the matter, but rather because doing it will result in something else that the person wants, such as a

reward. In this case, the true goal is to obtain the reward or another expected consequence, and completing the task is merely a means of doing so.

Many researchers have explored motivation theory in both the 20th and 21st century in ways that are specifically relevant to learning and education. Motivation theories as a collective group address both intrinsic and extrinsic forms of motivation in learning, and often in similar ways. Here is a quick review of some of the main motivation theories, as discussed by various sources (e.g. Kapp, 2012; Kim, Song, Lockee, & Burton, 2017) and their main findings:

- **Skinner's (1938) Operant Conditioning-**human behavior is a response conditioned by environmental stimuli; specific behaviors can be strengthened through the use of positive or negative reinforcement, or weakened through punishment; different types of reinforcement schedules (or reward schedules) can extend the motivational lifespan of a given behavior or drive it to extinction at different rates.
- **Keller's (1987-2009) ARCS Theory of Motivation-**in order to motivate a learner, one must successfully grab their *Attention*; offer content with high *Relevance* to the learners needs; provide a means of building their learning *Confidence*; and provide the learner with a sense of *Satisfaction* towards completing that learning task.
- **Elliot and Dweck's (1988) Achievement Goal Theory-**there are two types of goals: mastery and performance goals; learners with mastery goals focus on their personal learning achievements and intellectual growth; learners with performance goals focus on how their achievements compare to others.
- **Malone and Lepper's (1988) The Taxonomy of Intrinsic Motivations for Learning-**learning activities should provide goals that are challenging and whose progress can be gauged by the learner; the activity should arouse the learner's curiosity both through sensory stimuli and cognitive stimuli; learners should have a sense of autonomy and control throughout the learning process; fantasy can be applied to make learning more interesting cognitively and emotionally; learners can be motivated by different forms of social interaction, such as cooperation and competition; and learners like to be recognized for their performance.
- **Deci and Ryan's (2008) Self-Determination Theory**-in order to grow, learners: need autonomy or a sense of control; must feel competent enough to reach the goal; and have a sense of connectedness or belonging to a group.

2.2.2. Other Learning Theories Relevant to Gamification in Education

In addition to motivation theories, there are other theories that have contributed to gamified learning theory from behaviorist, cognitive, and constructivist learner approaches (Kapp, 2012; Kim et al., 2017).

- **Csíkszentmihályi's (1975) Flow-**a state of mind when one is fully mentally immersed in a given task; this state is usually marked by intense focus, the loss of track of time, and the learner's perceived ability to achieve a challenging task because its matches their perceived skill level.
- **Bruner's (1976) Scaffolding Theory-** tasks are assigned based on the learner's immediate capabilities and then progressively built onto with more difficult tasks; supporting learning through step tasks eventually allow the learner to obtain all the skills necessary to achieve the final goal.
- **Bandura's (1977) Social Learning Theory-**behavior is learned through social interactions based on simple observations of each other, imitating actions, and modeling.
- **Baddeley and Longman's (1978) Distributed Practice**-content learning is spaced out in multiple sessions, not crammed into one; spacing improves content retention.

By incorporating the main tenets of the above-mentioned theories, gamified learning is perceived as a valuable teaching approach because its implementation of game-based elements supports the kind of learning environment that naturally fosters the motivation of students to willingly participate and actively engage with the learning material. For example, reward structures respond to the need to build learner confidence and satisfaction, as well as operate as a form of social recognition of performance; goals and game levels respond to the need for creating a progressively challenging atmosphere while also breaking up the learning process into smaller achievable steps that support and develop the learner's total skill range; aesthetics and storytelling give way to fantasy, unlocking opportunities for extended learner curiosity; player modes such as competition and cooperation foster different kinds of social connectedness and allow students to both model and learn from each other's behavior; the curve of interest seeks to establish a flow in which learners can get "lost" in the experience; etc. Consequently, gamified learning is based on the assumption that the application of game-based elements to learning environments offers multiple and various strategic opportunities to optimize the learner's engagement in the learning activity, whether that is through appealing to learners'

intrinsic and extrinsic forms of motivation or attempting to channel other aspects of the learning process. In this way, the odds of engaging the learner are considered highly probable, and establishing successful engagement in the activity is often the precursor for more successful learning outcomes.

2.3. Early Game-Related Research Studies in Education

There is not a substantial reference base of studies that directly compare high tech and low tech gamification, which is likely due to the broad assumption that advanced technology is an automatic component of modern day gamification. However, there are similar studies in game-based learning that compared the use of high technology and games with traditional, non game-related teaching methods. While these studies are inherently distinct, they can be considered at least partially relevant to the topic of high tech gamification, and were arguably some of the stepping stones to gamified learning in general. Some researchers performed meta-analyses of the empirical studies conducted in this area, which revealed some key early findings of the positive impact of computer games and simulations (high tech innovations) in learning compared to traditionally formal instruction. Kapp (2012) did a review of these meta-analyses; the most relevant ones for this study and their findings are summarized below:

- **Randel (1992):** found that approximately a third of the studies on the effect of games and simulations in regular school subjects (e.g. language arts, biology and math), compared to traditional classroom instruction, had a positive effect on student performance, and that games in learning were likely to be more beneficial when applied to specific content and clear learning objectives.
- **Vogel (2006):** found that computer gaming and simulations overall revealed positive effects on students' cognitive gains and attitudes towards learning compared to traditional classroom teaching.
- **Ke (2009):** found that approximately half of the studies on computer-based instructional games had positive effects compared to conventional instruction.

2.4. Recent Studies in Gamified Learning

It was difficult to encounter peer reviewed empirical research studies pertaining to the exact scope of this particular study. However, there are existing studies that have suggested

overall positive outcomes as a result of gamified learning applied in related contexts. As a recall, the total amount of empirical studies on gamified learning in general are rather disproportionate to the total amount of papers discovered on the topic, according to the findings of several literature reviews/meta analyses. When narrowing in on empirical studies matching the variables of interest in this study, namely the EFL subject, CLIL context, and K-12 Spanish demographics, the results are quantifiably less impressive. Below is a summary of findings from the most recent and commonly referred to literature reviews and meta-analyses that can be applied to this study. The summary is divided into two sections: the first section outlines findings from reviews of studies exclusively about gamification and education; the second section outlines findings from reviews that broadly sampled all areas of gamification studies, but included relevant sub findings about gamified learning. The findings are listed in chronological order of the time frame the data was collected.

2.4.1. Literature Reviews on Gamified Learning

- **Caponetto et al. (2014):** sampled 119 theoretical and empirical studies between 2011-2014; only 7 came from researchers in Spain; only 3% of the studies represented primary school education and 4% for secondary school, compared to 43% at university level; the only mentioned study that focused on language was a theory based paper on Polish as a foreign language; the studies' outcomes on the effectiveness of gamification were not specified.
- **Dicheva et al. (2015):** sampled 34 empirical studies between 2011-mid-2014; only 2 papers targeted the K-12 education level, but none targeted languages as the main subject; majority results were reported as mostly positive.
- **Dichev and Dicheva (2017):** sampled 51 empirical studies between mid 2014-2015; only 7 papers targeted the K-12 education level, but none targeted languages as the main subject; of the studies on affective/behavioral/cognitive outcomes, only 12 reported positive results while 26 reported inconclusive results; the researchers cautioned against generalizing the results of the 12 positive studies based on the highly diverse factors of each study.
- **Majuri et al. (2018) / Koivisto & Hamari (2019):** sampled 128 papers up until mid 2015; the majority reported overall positive results, but there were also significant numbers for mixed outcomes.

- **Sailer & Homner (2019):** sampled 45 empirical quantitative studies up until March 2017; found small, but statistically significant effects on all cognitive, motivational, and behavioral outcomes.

2.4.2. Literature Reviews on Gamification Studies with Subfindings on Gamified Learning

- **Hamari et al. (2014):** sampled 24 empirical studies on gamification and motivational affordances between 2010-2013; 9 were specifically related to gamified education, and all of these were reported to have fully or at least partially positive learning outcomes.
- **Seaborn and Fels (2015):** sampled 30 theoretical and empirical studies on gamification in general from 2011-2013; only 8 empirical studies were applied in education, all targeted adult or higher education (undergraduate) populations, and 5 were reported positive.

In summary, surprisingly most of the literature reviews on gamification in education were conducted between 2011-2015, with only one study extending the sample time frame up to 2017. Upon further analysis, it was discovered that all of these studies used similar, yet distinct data collection procedures (e.g. data inclusion factors, search engines), which explains the variation in the number of studies sampled. In addition, they had similar but varied research goals, which affected the kinds of summaries produced about the samples collected, such as the regional origins of the study and the types of results. Based on these reviews, empirical studies that have been published on gamification and education have been largely reported as positive, or partially positive, which leaves enough reason to keep building upon the literature base and to encourage instructors/researchers to continue experimenting with how gamification works in the classroom. However, despite the positive rapport, it should not be overlooked that the quantifiable amount of empirical evidence for gamification in education as a whole is still extremely lacking, as pointed out by all reviews. This lack is further polarized when considering the evidence for gamification towards specific subcategories, such as gamification in *language* learning, gamification in the K-12 demographic range, and gamification in regional/country case studies. Due to these shortcomings discovered by the previously discussed literature, more empirical studies in gamification and all its subcategories should be openly invited. As such, this particular study aimed its research questions towards investigating more insights about

gamification in application to CLIL and EFL settings for 6th graders in Spain, a quest which not only contributes to the overarching literature on the topic of gamification in education, but also to neglected, albeit narrow, sub-paths including: languages as a subject, K-12 education demographics (particularly those of Gen Z), and regional case studies. Last but not least, this study intentionally aimed to bring more attention to the differences between the effective use of high tech and low tech gamification since the large majority of gamification literature assumed high technological features as a default setting.

3. Methodology

This chapter discusses the methodology used in this study in order to address the following research questions: 1) Does gamification result in improved EFL learning test scores for 6th grade Spanish CLIL students, regardless of the level of technology implemented (high tech/low tech)? and 2) Is high tech gamification more effective than low tech gamification in improving EFL test scores for 6th grade Spanish CLIL students?

3.1. Research Design

In this research study, a quantitative research approach with a case study design was used to measure and compare how well Spanish CLIL EFL students performed on improving their language vocabulary skills through two types of gamified learning: high tech gamification and low tech gamification. The participants of this study were randomly divided into two groups, an experimental group and a control group. The experimental group took part in high tech gamified activities via a gamified learning website called Quizlet, while the control group took part in low tech gamified activities that were adapted versions of the Quizlet activities. Students' academic performance was measured by comparing test scores taken from a pre test and post test, which were analyzed using descriptive statistics and then tested for statistical significance with corresponding *t*-tests.

3.1.1. Class Intervention

Aside from being the source of data for this research study, this experiment was a real-life, meaningful school project. The goal of this short-term project was to help the students practice and improve their basic English vocabulary knowledge, as it pertained to their age group and level as A2[2] level learners. This was considered an important and relevant teaching and learning objective, not only for the students' likely ongoing future need of basic English vocabulary skills, but also as a valuable opportunity to prepare them for their upcoming end-of-the-year Cambridge English: KEY (KET)/ A2 KEY examination. Students' scores on this exam would help qualify their access to continued English CLIL-based education in an upper secondary school. In light of this context, the study was designed to be an extracurricular called

[2] In the Common European Framework of Reference for Languages

English Club, which was offered and set up during after school hours as a completely free and voluntary English language club for students only of the sixth grade. Students were informed that participating in the club/study would in no way affect their real academic performance. There was no punishment if students decided to drop the club at any given time and attendance was completely based on their own free will.

The instruction schedule was set for twice a week 30-min after-school sessions, which were offered on a fixed block schedule; the experimental group was assigned Mondays and Wednesdays in the computer lab, and the control group, Tuesdays and Thursdays in the classroom. This set-up was influenced by external factors: first, the researcher's availability to work with the participants was limited to a four-day work week (as opposed to a 5-day week), as decided by an existing work contract; and second, although the researcher initially desired to conduct longer sessions (ideally an hour), the length of the sessions was inevitably determined by the closing hours of the school campus as well as the fixed schedule of the after school lunch program in which many of the participants were part of. In addition, due to these constraints, lessons were held immediately after the final teaching hour, which ultimately resulted in some lost time. For instance, although sessions always ended on time, they did not always begin punctually due to the trickle-in attendance of students transitioning from their last period; the researcher's estimate of the true average length of each session would be approximately 25 minutes.

The learning content was taught in a distributed practice style that was delivered to both groups in the same, randomized order, and in cumulative fashion. Thus over the course of the study, new content was introduced and old content was being reinforced simultaneously. The study lasted the length of a 9 week period, in which 21 sessions were successfully conducted, 11 sessions were for the control group, and 10 sessions for the experimental group, excluding the pre and post test sessions. More completed sessions for each group were anticipated, but due to the national/regional Spanish academic school calendar and the private calendar set forth by the school itself, the study could not be conducted on an uninterrupted consecutive weekly basis. In fact, during the 9 week period, the study was forced into 4 temporary pauses, accounting for the loss of 11 potential sessions, in respect of student/teacher vacations/holidays and other unforeseen school-related scheduled priorities. However only data based on the learning content from the first 10 sessions of each group were included in the final pre and post test marking procedures in order to maintain a fair balance in the data collection and analysis.

3.2. Experimental Variable

The experimental variable in this study was the application of high tech gamification, as a complementary teaching method in a CLIL EFL context. The technology specifically chosen to represent high tech gamification in this study was a gamified learning website online called Quizlet. This website was chosen as the representative example of high tech gamification for practical and convenience-based reasons: 1) it is a free resource available to the general public; 2) it is easy for teachers to join with a simple email-based account; 3) it is not obligatory for students to create a private account in order to participate in online activities, thereby protecting the privacy of minors; 4) basic usage of the games in class requires relatively standardized, high tech elements, namely a computer and a reliable Internet connection; a projector and projector screen are required for optional group activities and; 5) the site interface is relatively simple and straightforward to navigate for all users.

Fortunately, the school where this study took place had a small computer lab with mostly reliable computer hardware, a stable Internet connection, and a classroom projector and screen. Consequently, Quizlet was an easily accessible online gamified learning tool to exemplify how high tech gamification could be implemented into EFL CLIL lessons when basic technological resources, like the ones mentioned above, are readily available. Despite the choice of Quizlet as the main gamified learning tool, this study was not designed to emphasize the value of Quizlet over other gamified learning websites, nor over other forms of high tech gamification.

3.3. Control Variables

Both the experimental and control groups were shaped by three main control variables:

1) English CLIL-based instruction
2) identical language learning content
3) gamified learning activities

Of these controlled variables, the first was considered a long-term default setting. The latter two required more careful attention during the research design process, which will be explained in the following sections.

3.3.1. Control Variable 1: Engligh CLIL-Based Instruction

English as the target foreign language was used as the sole instructive language throughout the entire study. The instructor/researcher administering the study priorly possessed the formal educational role of the head native English language assistant for all of the participants in the study. As part of the requirements of being a native language assistant hired as part of the Bilingual and Bicultural Project, providing full language immersion while in contact with all CLIL EFL students was a serious priority in the classroom. Based on this reality, incidents of accidental communication between the researcher and the participants in their native language, Spanish, was never reported as an issue during the time of the study.

3.3.2. Control Variable 2: Language Learning Content

One of the main advantages of using vocabulary learning content was that it provided a clear, quantifiable EFL learning outcome that was relevant towards the first research question, which sought to discover if gamification resulted in improved scores despite being characteristically high tech or low tech. Another advantage was that the test tasks solicited the acquired vocabulary knowledge in ways that were expected of/familiar to the participants. This reduced the risk that students might lose points for failing to understand the task objectives, which would have risked the accuracy of the scores.

A2 level of English vocabulary terms were considered a practical content choice since students were expected to encounter these terms throughout the A2 KEY examination. Terms were selected from an official A2 KEY vocabulary list, which was directly accessed from the Cambridge English website as a free preparation resource for teachers and students. The list was published in 2012 by the British Council and the University of Cambridge Local Examination Syndicate (UCLES), also known as Cambridge English Language Assessment/Cambridge Assessment, which are the groups responsible for organizing and administering a wide range of English qualification exams at schools internationally. The original A2 KEY vocabulary list comprised of 25 different vocabulary topics. In order to narrow down the selection due to real time constraints, 125 vocabulary terms were initially randomly selected from the list with an online random choice generator tool (textfixer.com). The selected terms were then regrouped into categories for practical teaching purposes. By the end of the term, there proved only sufficient time to introduce 75 vocabulary terms out of the original 125.

3.3.3. Control Variable 3: Gamified Learning Activities

Beyond the researcher's curiosity to explore high and low tech gamification, using gamification as a controlled variable in both the experimental and control groups was a strategy to effectively recruit enough participants. In other words, the researcher expected that attempting to recruit participants for a standard, non-gamified English Club would ultimately result in little participant interest to join the club, the misconstrual of it as a traditional lesson/book based study group, and/or a rapid participant mortality rate, thus compromising the study's data.

Quizlet, as the chosen source of gamified learning activities, offered several different study and play features, but due to the fixed time constraints of each session, not all of these features could be utilized during the study. In order to keep participants engaged in the gamified learning process without being overwhelmed by the usage of superfluous amounts of new learning activities, the study was designed around four different core learning activities directly from Quizlet. While these original Quizlet activities were used in the experimental group, the control group used parallel activities that were based on the same Quizlet activities.

In order to accomplish this, the conceptual game-based elements and basic activity designs used throughout the gamified activities in Quizlet had to be isolated from their high technological context. Extracting these designs and elements allowed the activities to be reproduced as similarly as possible in the low tech design of the gamified activities of the control group. In short, the gamified activities used in Quizlet by the experimental group were adapted as necessary into low tech versions for the control group. The intent was to reduce extreme differentiation between the gamified activities of each group, thus isolating the experimental variable, high technology, as much as possible.

The four activities were instructed and fulfilled in a guided, logical step-by-step pattern: 1) learning new content via flashcards 2) reinforcing/mastering the content with a multiple choice style activity, 3) competing for individual high scores in a matching game and 4) competing in a team racing based game. See Table 2 for details.

Table 2. Comparing Gamified Activities as a Control Variable

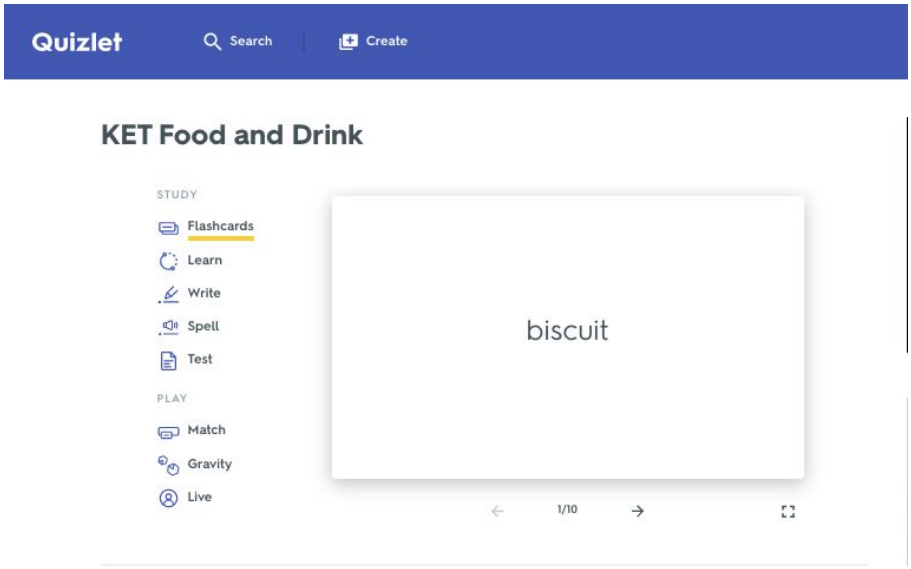
TYPE OF ACTIVITY & GAME-BASED ELEMENTS INVOLVED	EXPERIMENTAL GROUP Quizlet's original high tech gamified activities	CONTROL GROUP Adapted low tech gamified activities
1. FLASHCARDS -cooperation mode -time based -goal/level 1 scaffolding -replay/do over -curve of interest (entry point) See Figure 1.	MATERIALS Computer with Internet access, Quizlet's "Flashcards" study features	MATERIALS paper flashcards
	RULES In pairs, students use a digital deck of flashcards to study terms and definitions.	RULES In pairs, students use a paper deck of flashcards to study the terms and definitions.
2. MULTIPLE CHOICE -competition mode -time based -goal/level 2 scaffolding -feedback -replay/do over -curve of interest (challenge 1: test your knowledge) See Figure 2.	MATERIALS Computer with Internet access, Quizlet's "Learn" study feature	MATERIALS paper flashcards
	RULES Students individually take turns to read randomly generated definitions and select the correct corresponding term from the choices given on the screen. Answers are automatically checked after every card.	RULES Every student individually spreads out a deck of cards. The instructor reads out definitions in a randomized order while students pick up the corresponding card. Cards are flashed to the teacher for accuracy checking.
3. MATCHING -competition or conflict mode -time based -goal/level 3 scaffolding -feedback -reward structures: high scores, points and leaderboards -replay/do over -curve of interest (challenge 2: compete with your knowledge) See Figure 3.	MATERIALS Computer with Internet access, Quizlet's "Match" play feature	MATERIALS paper flashcards
	RULES Students take turns individually matching definitions and vocabulary words scattered on the screen with a built-in timer feature. Time records are reported to and documented by the instructor. The lowest (fastest) scores are reflected on the teacher's manual group scoreboard.	RULES In pairs, students spread out one deck of flashcards. The instructor calls out a prompt. Students compete one-on-one to pick up the correct corresponding word/definition the fastest. Every card has a one-point value. Points are reported to and documented by the instructor. The highest scores

		are reflected on the teacher's manual group scoreboard.
4. TEAM QUIZ RACING -cooperation and competition mode -time based -goal/level 4 (use all acquired skills) -feedback -reward structures: leaderboards, points -replay/do over -curve of interest (climax) See Figure 4.	MATERIALS Computer with Internet access, projector and screen, Quizlet's "Live" play feature	MATERIALS Flyswatters, A4 paper sized vocabulary cards, adhesive, flat wall surface
	RULES Teams are randomly generated. Computers belonging to the same team will display the same vocabulary prompts, but each with a different set of possible answers. Only one team member's computer has the correct answer, forcing each team to work in collaboration. Wrong answers are penalized by the team's score getting set back to zero to start over. Teams race to reach the winning point level. Team progress is openly displayed throughout the game on the projector screen.	RULES Vocabulary cards are randomly fixed to a flat wall surface (e.g. the blackboard). The instructor randomly divides students into teams. The instructor calls out vocabulary prompts and students race one-on-one to touch the correct card with the team flyswatter. Teams collaborate freely, but must take turns using the flyswatter. Wrong answers are penalized by a point deduction. Teams race to reach the winning point level. Team progress is openly displayed on the board with a simple tally system.

Table 2 describes the high and low tech gamified activities used in the experimental and control groups with information regarding: 1) the type of activity and the main game-based elements involved in that activity, and 2) the materials used and the basic rules of each gamified activity per group. There is a distinction made between the experimental group activities shown in column 2, which were original Quizlet gamified activities, and the control group activities shown in column 3, which were adaptations of the Quizlet activities in a low tech environment. Figures 1-4 were also provided to offer complementary visual representations of the Quizlet activities to show how the activities functioned during the study, however visual examples of the use of the adapted low tech versions live in the classroom were not collected to protect participant privacy.

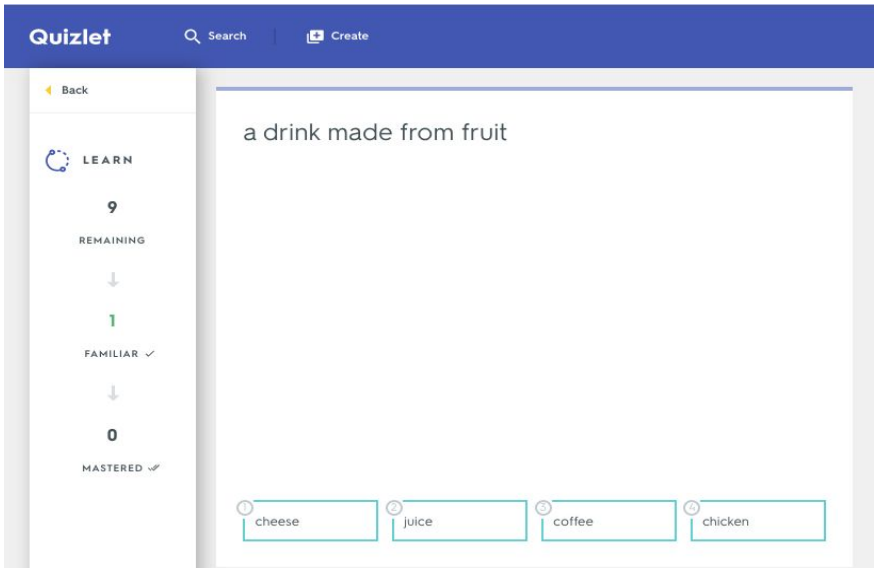
Figures 1-4. Quizlet's Gamified Learning Features

Figure 1. Activity 1: Flashcards



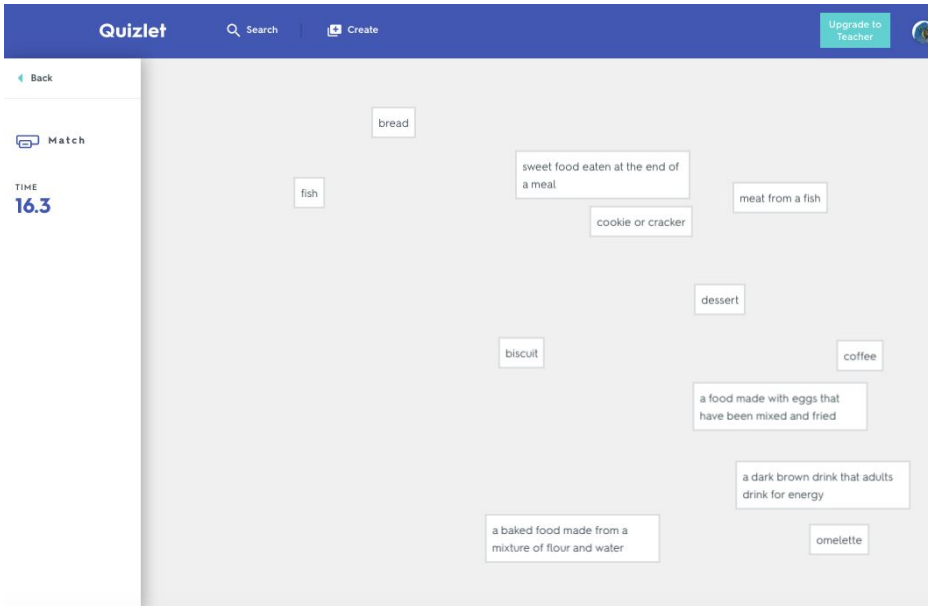
Students quickly flipped through a digital set of flashcards on the computer screen as the first step to learning content.

Figure 2. Activity 2: Learn



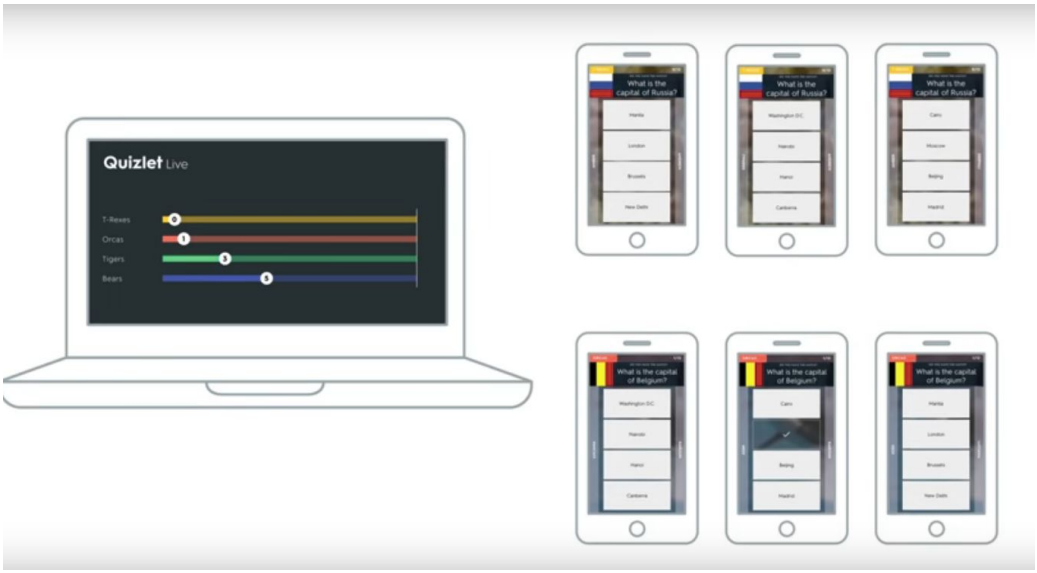
Students reinforced the content by playing a multiple choice quiz with instant feedback and a progress sidebar.

Figure 3. Activity 3: Match



Students dragged and dropped matching terms and definitions as quickly as possible using the time feature to gauge their progress.

Figure 4. Activity 4: Live



Students were randomly grouped into computer teams and worked in collaboration to correctly answer all of their prompts before the competing team; the game was initiated and controlled by the instructor's computer, and team progress was automatically displayed on the projector screen.

Figures 1-3 were screenshots recorded to demonstrate real glimpses of the first three activities used in this study; they feature examples of the actual learning content used in the study. Figure 4 is only a similar representation of the general set-up of Activity 4: Live. While a laptop and tablets/smart phones are displayed in the image, in reality they were not available for neither the instructor nor the students; in the actual study, desktop computers replaced mobile devices.

In review, the four Quizlet activity features selected for the gamified research design were Flashcards (Figure 1), Learn (Figure 2), Match (Figure 3) and Live (Figure 4), as explained in Table 2. All features were selected based on their ease of use in respect to the available time per session. Quizlet's other features, Write, Spell, Test and Gravity, were disregarded for the following reasons: Write, Spell and Gravity required basic typing proficiency skills that the students did not possess at the time of the study and there would not have been enough time to learn; Spell required audio devices that were not reliably functional; and Test was dismissed as a standard testing activity because its main activity was designed in an uncreative multiple choice test task format. Given that the activities were adapted into low tech gamified versions for the control group, identically matching the activities between both groups was not always feasible or possible, so certain changes were implemented logically. The only activity in which the adaptation was straightforward was Flashcards. In Multiple Choice for the control group, multiple choice exams were opted out of to avoid implementing an activity based on a traditional test format. In Matching for the control group, individual timers were not readily available to mirror the use of the built-in timer feature in the experimental group. To account for this, correct card answers were given point values in order to foster the competitive atmosphere originally induced by the timer effect. In the experimental group, Quizlet scores could have been automatically tracked and saved online instead of logged manually by the instructor, but it was avoided because the feature required student logins. Lastly, Live in the control group was adapted in an attempt to accommodate the large cumulative growth of learning content, which would have been too chaotic to physically reproduce and redistribute to participants during the limited time frame of each session. It is acknowledged that the mirroring process of the gamified activity designs from high tech to low tech was clearly not 100% exact. Still, all adaptations were designed to reflect the same/similar kind and number of game-based elements used in the original Quizlet versions as closely as possible, with the overall intent to keep all controlled variables constant.

3.4. Participants

The participants in this study were selected from a convenience-based sample. The researcher/administrator of this study was their native English language assistant during their English CLIL classes for the academic year of 2018-2019, so they were readily available test subjects. At the time of the study, the participants were native Spanish speaking students in the 6th grade, aged 11-12. They were in their final year of primary education and represented the entire 6th grade at the school.

3.4.1. School Background

The participants of this study came from a public bilingual primary school that is part of the Spanish Ministry of Education, Culture and Sports' Bilingual and Bicultural Project, and currently offers a CLIL EFL based curriculum to its students. Of the two streams offered by bilingual centers, this school offered "Sección," which entails that the students are taught a minimum of three courses in English, in comparison to "Programa", in which the minimum would have been one course in English (Shepherd and Ainsworth, 2017). The school is located within the autonomous Community of Madrid in Spain. Madrid is an especially active region for CLIL, where by 2015-2016, the bilingual project had already been installed into well over 400 primary and secondary schools total. Although Madrid as a region technically ranks at a high EFL proficiency level, it still falls a solid 12-15 points below the top ranking non-native English speaking European nations (Shepherd and Ainsworth, 2017), and thus was coincidentally an interesting region for the sample pool.

3.4.2. Study Sample

The original sample pool consisted of 33 participants, who originally came from two separate classes. Students from both classes were conglomerated into one sampling pool, the reason being their equal background status in age, grade level, and school, and then randomly assigned to the experimental group, Group A, or the control group, Group B. Two were automatically disqualified from the data sample as outliers due to previously existing learning disabilities/disadvantages (these participants were still welcomed to participate, but ultimately did not express interest). Of the remaining 31 students, 26 students freely volunteered to

participate. By the end of the study, 21 students qualified for the final data sample; 5 were disqualified due to lack of attendance/early quitting.

Each group was intended to have the same amount of participants, however, due to special student circumstances (e.g. colliding extracurricular schedules), some participants were obligated to switch groups after the initial random distribution. Participants were unaware at the time of sign-up that there would be two distinct group types. After the finalized group assignments were released, requests to change individual placements were not permitted due to fairness procedures, the popularity of the club, and the limited availability of fully functioning computers for the experimental group.

The experimental group, Group A, originally consisted of 15 original participants, which reduced to 13 by the end of the study, due to excessive absences/loss of interest in the club. The control group, Group B, originally consisted of 11 original participants, which reduced to 8 by the end of the study for the same reasons.

3.4.3. Ethical Issues

Given that there was no perceived personal danger or threat to any of the participants involved in the club (perhaps aside from the eardrum trauma of the researcher), oral consent was deemed sufficient from participants and their parents to join the club, followed by a written signature on the sign up list. Permission to conduct this after school extracurricular club was welcomed by the authoritative school staff and monitored by the same staff throughout the entire study. Approval to use data from the club was given under the condition of participant and school anonymity. In respect of this condition, no personally identifying information from any of the participants or the school was used in this report.

3.5. Data Collection

3.5.1. Instruments: Pre and Post Tests

Since the goal was to provide the participants with general exposure to and reinforce as many vocabulary terms before the A2 KET exam, the pre and post tests consisted of basic types of questions. Teaching the vocabulary content at a greater knowledge depth would have been a slower and lengthier process that such a short study would not have afforded, especially with the soon approaching KET exam deadline. The objectives of the test questions were to simply measure if a student could recall the basic meaning of a term, be able to match it or

produce the correct term in accordance to the prompt, and be able to spell the term correctly if required. Test questions included activities that required matching terms to definitions, categorizing terms in groups, matching terms to simple graphics, answering true or false questions, and completing word production prompts. These kinds of questions were chosen because they reflected those regularly found in the participants' regular A2 English language course textbooks and class assignments given throughout their regular CLIL EFL classes.

The pre and post tests were designed to contain the same types of test questions and the same vocabulary learning content, but varied by total score. The pre test assessed the original 125 vocabulary terms while the post test ultimately assessed only 43 of these terms. This is due to the fact that at the beginning of the study, the pre test was designed to test how many terms the students already understood prior to being exposed to the treatment, gamified learning activities, and the researcher had estimated that there would have been enough time in the schedule to introduce all of the targeted content. However, in retrospect, this was a highly optimistic goal, and unfortunately, there was only time to introduce 75 of these terms during the length of the study. Consequently, it would have been illogical to test for all of the same terms in the post test, since much of it was not covered. Ultimately, the researcher designed the post test around 43 terms that were randomly selected from only the content taught during the gamified lessons to each group. This final test content is displayed in the Appendix.

3.5.1.1. Marking Procedure

In order to account for the mismatching total scores of the pre and post tests, any terms that were not introduced at all during the study were excluded from the marking procedure of the pretest in order to only include data from content that was actually studied. In addition, since there were only 43 vocabulary terms assessed in the post test, only the questions/content of the pretest matching these terms were marked. This procedure ensured that the pre test and post test had the same final raw totals of 43 points measuring the same content. For time-effective and objective grading purposes, all tests questions were assigned a 1-point value, which either received full or zero credit. When applicable, answers had to be spelled correctly to receive credit.

3.5.1.2. Distribution of Tests

The same pre and post tests were administered to each group in the classroom before and after the gamified study, respectively. No more than 30 minutes was allotted for participants to finish the tests, but most participants finished early. Participants were not allowed to ask for help during the exam and were encouraged to simply try their best.

3.6. Analysis Procedures

Testing the results for statistical significance meant providing more meaningful scientific evidence that was useful towards the second research question, which aimed to investigate the effectiveness of high tech gamification over low tech gamification in learning. This was seen as a more valuable method of analysis than merely observing the data and loosely drawing conclusions or suggestions about whether high/low tech gamification would be more effective than the other.

3.6.1. Raw Data

The raw scores were taken from the pre and post test data samples from both the control and experimental group. The raw data was used to find the mean score and standard deviation of each sample. These scores were also shown in their equivalent percentages for discussion purposes, but only the raw scores were used in the actual data analysis. See Tables 3 and 4 in Chapter 4 Results.

3.6.2. Outliers

Potential outliers in the data samples were checked by the Interquartile Range (IQR) method using Alcula's (2019) IQR calculator. According to the IQR rule, 1.5 x (IQR) is subtracted from the first quartile and added to the third quartile, determining the minimum and maximum intervals for all normal data points. Any data points outside of this range are considered outliers. As a second measure, outliers were visually checked with box plots generated by Alcula's (2019) box plot calculator. See Table 5 and Figures 5-8 in Chapter 4 Results.

3.6.3. Tests of Equal Variances and Normality

Levene's test was used to check if the data samples of the pre and post tests of both the control and experimental group had equal variances and was calculated with the Homogeneity of Variance calculator by Stangroom, (n.d.).

The Shapiro-Wilk W-test was also used to check if the distribution of the data samples of each group was normal. This test was chosen because the sample sizes of each group were less than 50, and was calculated with the Shapiro-Wilk calculator by Statistics Kingdom (n.d.).

3.6.4. Parametric and Nonparametric *t*-tests

Either a student *t*-test or the Wilcoxon signed-rank test was used to discern if there was any statistically significant improvement between the pre and post tests of the control group, as well as the experimental group. Based on the results of Levene's test and Shapiro-Wilk's test, if the data samples of a group reported to have equal variances and a normal distribution, the parametric student *t*-test was used; if the data did not have both equal variances and a normal distribution, the nonparametric equivalent *t*-test, the Wilcoxon test, was used. These tests were calculated with the *t*-test calculator by GraphPad (2018) and the Wilcoxon signed-rank calculator by Statistics Kingdom (n.d.). See Table 6 and 7 in Chapter 4 Results.

4.Results

4.1. Raw Data Results

The results of the marking procedure used on the pre and post test scores of each group are displayed in Table 3 and then further described in Table 4.

Table 3. Raw Data and Percentages

	Control Group				Experimental Group			
Participant	**Pre test scores**		**Post test scores**		**Pre test scores**		**Post test scores**	
1	**22**	*51.16%*	↑**28**	*65.12%*	**37**	*86.05%*	↑**41**	*95.35%*
2	**38**	*88.37%*	↓**37**	*86.01%*	**36**	*83.72%*	↑**39**	*90.70%*
3	**27**	*62.79%*	↑**36**	*83.72%*	**40**	*93.02%*	↑**41**	*95.35%*
4	**36**	*83.72%*	↓**31**	*72.09%*	**41**	*95.35%*	↓**40**	*93.02%*
5	**27**	*62.79%*	↑**37**	*86.05%*	**33**	*76.74%*	↑**34**	*79.07%*
6	**26**	*60.47%*	↑**28**	*65.12%*	**40**	*93.02%*	↑**42**	*97.67%*
7	**38**	*88.37%*	↑**40**	*93.02%*	**32**	*74.42%*	↑**36**	*83.72%*
8	**27**	*62.79%*	↑**35**	*81.40%*	**30**	*69.77%*	↓**29**	*67.44%*
9					**39**	*90.70%*	↑**42**	*97.67%*
10					**34**	*79.07%*	↑**40**	*93.02%*
11					**29**	*67.44%*	↑**32**	*74.42%*
12					**39**	*90.70%*	↑**41**	*93.35%*
13					**32**	*74.42%*	↑**34**	*79.07%*

Table 3 reflects the pre and post test raw scores collected from each participant in each data sample. The total raw scores for both the pre and post tests were out of a total of 43 points. The up and down arrows symbolize if the raw post test scores increased or decreased.

Table 4. Means and Standard Deviations

	Control Group n = 8		Experimental Group n = 13	
	Pre test	**Post test**	**Pre test**	**Post test**
mean	30.13	34	35.54	37.77
SD	6.22	4.47	4.12	4.28

Table 4 reflects the means, or averages of the scores from the data given in Table 3. The standard deviation is represented by "*SD"* while the population size of each sample group is represented by "n."

4.2. IQR and Boxplot Outlier Analysis Results

The results of performing the IQR and boxplot outlier methods on the data are displayed in Table 5 and Figures 5-8.

Table 5. Descriptive IQR Outlier Statistics

	Control Group		Experimental Group	
	Pre test	**Post test**	**Pre test**	**Post Test**
Quartile 1(Q1)	26.25	28.75	32	34
Quartile 3 (Q3)	37.5	37	39.5	41
IQR (Q3 - Q1)	11.25	8.25	7.5	7
Minimum and maximum interval	9.38-54.38	16.38-49.38	20.75-50.75	23.5-51.5

According to the descriptive statistics shown in Table 5, none of the data points fell outside of the minimum and maximum interval range, so no significant outliers were detected.

Figures 5-8. Box Plot Outlier Results

Figure 5. Control Group Pre test

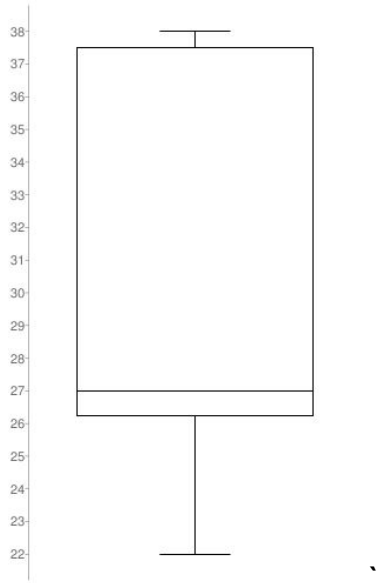


Figure 6. Control Group Post test

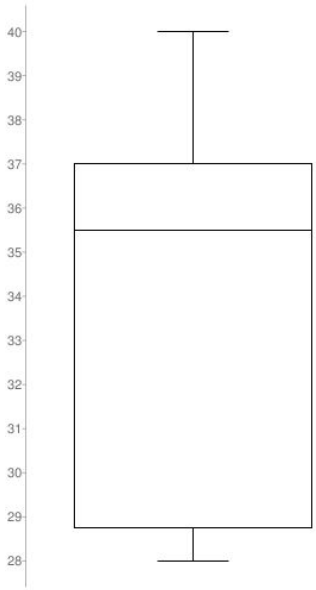


Figure 7. Experimental Group Pre test

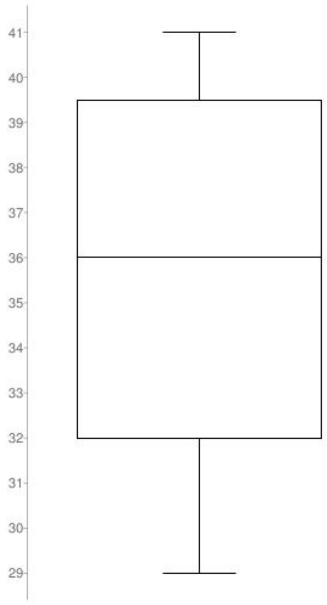
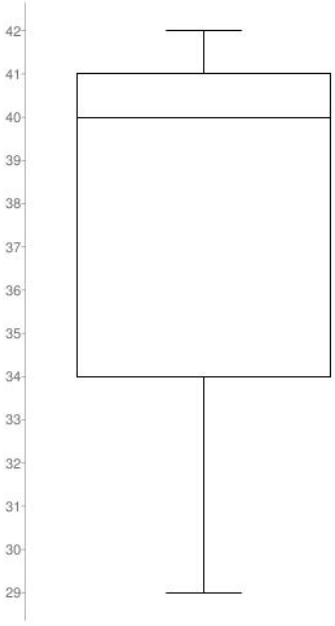


Figure 8. Experimental Group Post test



Outliers in box plots are normally represented by points marked outside of the minimum and maximum values. According to Figures 5 and 6 of the control group and Figures 7 and 8 of the experimental group, there are no visually detected outliers displayed in any of the box plots of each group's respective pre and post test data. These results supported the IQR outlier results.

4.3. Levene's Test and Shapiro-Wilk's Test Results

In respect to the data samples of the control group, the results of Levene's test showed an *f*-ratio value of 2.38 and a *p*-value of 0.14. The same test on the sample data of the experimental group showed an *f*-ratio of 0.01694 and a *p*-value of 0.897516. These values signify that the results were not significant at the p<.05 level. Thus, both group samples met the requirement of equal variances.

The results of the Shapiro-Wilk's test showed *p*-values of 0.0772499 and 0.299356 for the control group's pre and post test data, respectively. This signifies that this group's sample data was normally distributed. The experimental group's results showed *p*-values of 0.281456 and 0.040595 for its pre and post test data, respectively. Contrary to the control group's results, these values signify that *only* the sample data of the experimental group's pre test was normally distributed.

Since Levene's test and Shapiro-Wilk's test results from the pre and post test samples of the control group both reported normal, this group's data qualified to be tested for statistical significance using the student *t*-test. On the other hand, since the results of the experimental group's post test sample failed the normality test, the Wilcoxon signed-rank test was chosen as a better fit for testing this group's data for statistical significance.

4.4. Student *t*-Test and Wilcoxon Signed-Rank Test Results

The results of the student *t*-test performed on the control group's data and the Wilcoxon signed-rank test on the experimental group's data are displayed in Tables 6 and 7.

Table 6. Student *t*-test - Control Group

n	df	*t*-value	Critical *t*-value	Standard error of difference	Two tailed *p*-value	95% confidence interval
8	7	2.0772	2.365	1.865	0.0764	From 8.29 to 0.54

Based on the student *t*-test performed on the control group's data, the null hypothesis was that there would be no difference between the pre and post test scores after the

implementation of low tech gamified learning activities. According to Table 6 results, the *t*-value was less than the critical *t*-value at 7 degrees of freedom. In addition, the *p*-value was not considered small enough at .0764. At a significance level of $p<0.05$, this means that the null hypothesis was accepted; the difference between pre and post test scores was *not* found to be statistically significant.

Table 7. Wilcoxon Signed-Rank Test- Experimental Group

n	Z statistic 95% / Critical value accepted range	Two-tailed *p*-value	W statistic / 95% accepted range
13	-2.993336 / [-1.96 : 1.96]	0.00275946	5 / [17 : 74]

Based on the Wilcoxon signed-rank test performed on the experimental group's data, the null hypothesis was that there would be no difference between the pre and post test scores after the implementation of high tech gamified learning activities. According to Table 7, the Z and W statistics fell outside of the accepted value ranges; however, since n=13 and was less than 25, critical values were used instead. The *p*-value was small at $p=0.0028$. At a significance level of $p<0.05$, this means that the null hypothesis was rejected and the alternate hypothesis was accepted; the difference between the pre and post test scores was found to be statistically significant.

5. Discussion of Results

The final results of the tests performed on the data samples of both the control and experimental groups were discussed from statistical perspectives and practical understandings. The discussions of these results and their implications were organized per research question:

5.1. Research Question 1

Does gamification result in improved EFL learning test scores for 6th grade Spanish CLIL students, regardless of the level of technology implemented (high tech/low tech)?

The researcher hypothesized that both high and low tech forms of gamification would improve the EFL learning test scores of the participants in this study. Based on the raw scores of each group, there was evidence that both forms of gamification on average did actually improve test scores, so the hypothesis that gamification would have a positive impact on test scores was overall correct. Upon closer inspection of the data, the results showed that of the 21 total participants, 19 performed better on their post test after being exposed to gamification, while only 4 participants did worse. Proportionally speaking, this means that low and high tech gamification had a positive impact on the test scores of 90% of all participants. By individual group assessment, 75% of participants in the control group performed better on the post test compared to 84.62% of the experimental group. These figures provided a comparative visual demonstration of the extent of the ability of each kind of gamification to influence learner growth, both which resulted in a positive growth for the strong majority of participants. These visibly enhanced outcomes may be regarded as supporting evidence of the potential success for gamification as a valuable teaching method.

Analyzing the point value differences in both groups, it was found that participants scored from a minimum of 1 to a maximum of 10 points higher in their post tests. In the context of the test size, any incremental gain could have a big impact on a student's marks. For instance, ten points was the difference between a barely sufficient passing score (in Spanish primary education, "*aprobado"*) and a very good/remarkable passing score ("notable") for participant #3 of the control group. For participant #9 of the experimental group, three extra points ranked the student's score higher up the same ranking bracket, turning a lower 90.7% score to an upper 97.67%. In the case of participant #4 of the control group, missing just one

point meant the difference between a good and a very good score, which although sounds similar, were actually two distinct measures of academic success. In all of these cases, whether the change was considered a drastic or relatively minor improvement, every point mattered. Most teachers would probably agree that by the 6th grade of primary school (if not earlier), students are well conditioned by the unspoken belief that success is measured by academic performance, which in turn is determined by points and grades. Thus, from a student's perspective, there is often a significant meaning given to a test score that upgrades his/her academic reputation from being a mere "average" student to an "outstanding" student, or that changes a 98% test score to a 100%. From a macro perspective, sometimes, minor improvements are enough to help a struggling student achieve the target goal of simply passing a course and continuing to the next grade level, and other times these improvements may help a student pass a special exam like the KET whose final test score could actually dictate the future direction of that student's academic career. Minor score gains such as the ones found in this small study ultimately do have real consequences not only on students' immediate self-perceptions of their intellectual growth, individual pride and motivation to keep learning and improving, but also on their ongoing educational journey into early adulthood, and for these reasons, gamification should be taken as a serious approach in the classroom.

In addition, another reason to consider gamification is that it is a flexible method that can be adapted to most teaching environments with or without reliable access to high technology. This perspective on gamification in education is useful, particularly for lower income schools for which low tech, and even no tech, classrooms and resources may still be the dominant norm. Since it should never be assumed that students in today's 21st century schools have ready access to high technology resources, it is ideal knowing that gamification can be adjusted to any given technological/economic school situation. In sum, the results of this study found that gamification can be a potential driver of enhanced student learning and improvement, so it deserves to be recognized and considered as one possible alternative solution to help improve the EFL learning experience in Spanish CLIL elementary contexts.

5.2. Research Question 2

Is high tech gamification more effective than low tech gamification in improving EFL test scores for 6th grade Spanish CLIL students?

The researcher hypothesized that high tech gamification would indeed be more effective in improving test scores than low tech gamification; more specifically, the researcher predicted that the increase in test scores of the high tech gamified group would be statistically significant while that of the low tech gamified group would not. According to the results of the student *t*-test of the control group's data, the *p*-value was .0264 points *greater than* the p-significance level of 0.05, while in the Wilcoxon signed-rank test of the experimental group, the *p*-value was .04724 points *less than* the same significance level. These values thereby allowed the alternative hypothesis to be accepted and the null hypothesis to be rejected; high tech gamification resulted in a statistically significant difference between the pre and post tests of the experimental group, but not for the control group. In other words, the use of high tech gamification was more effective in increasing test scores than the use of low tech gamification.

These results were interpreted in several useful ways. To begin, they have helped provide the basic grounds for researchers to purposely distinguish between two different forms of gamification based on their effective strength of improving test scores. In a field in which the existing literature has shown bias towards mostly high tech gamification studies, it is important for future researchers to begin more clearly identifying which types of gamification they use, given that high and low tech gamification can not be assumed to possess the same level of effectiveness. Since past studies have rarely acknowledged the concept of low tech gamification, it is possible that scholars in the field have been careless with their frequent conflation of the descriptive term "modern" with the term "high tech" when describing and using gamification in their studies. However, modernity and technological advancement are not inherently synonymous concepts, and actually both high tech *and* low tech gamification can be considered two different forms of modern teaching methods. The lack of statistical significance backing low tech gamification does not render it as completely futile as a teaching method, but rather shows that it is a topic area that remains largely underdeveloped. There are simply not enough studies to suggest or generalize the idea that low tech gamification may not possibly produce significant results in other learning situations, and it should not be forgotten that low

tech gamification had a positive impact on test scores. Instead, the researcher would argue that any method that results in any form of observable enhanced performance should continue being explored by researchers and instructors as an alternative teaching method, and given the struggling status of EFL proficiency on a nationwide level in Spain, any progress, statistically significant or otherwise, is still progress.

In continuation, a clear distinction between low and high gamification based on effectiveness makes it easier for instructors who have the lucky privilege to choose between the two forms, as determined by the technological resources available for students at their schools. For example, in the case of the participants/school involved in this study, it would be more productive for instructors to design future gamified lessons through the use of the existing available technology on campus, rather than through the use of traditional classroom materials. This is valuable information, not only because high tech gamification proved to be more effective than low tech, but also because of the observed comparative relative ease of implementing it. For example, it was noticeable how much more time-effective it was for participants to perform high tech gamified activities vs low tech gamified activities; in the high tech gamified group, participants simply sat down at a computer with the materials and activities lessons immediately available on their computer screens with little instructor interference, but in the low tech group, materials constantly had to be physically distributed and set up for each activity. Consequently, there was more time for participants to fully interact in the computer gamified activities, which allowed them to better maximize the already short learning sessions. This was not the case for the low tech group, whose gamified activities seemed to be overall more rushed and often also more stressful for the instructor. Moreover, participants of the high tech group seemed to be generally more excited about coming to English Club compared to the low tech group, whose members had inquired as to why they could not switch groups. Logically speaking, it is easier to teach new content when the students enjoy the activities involved, and since most participants expressed preference or contentedness for the group with high tech gamification over low tech gamification (either directly by voicing their opinions or indirectly through (in)consistent attendance), it would be foolish for instructors to choose the latter when the two gamification options exist.

Lastly, another interpretation of the results was that they provided reason to support the researcher's additional goal to investigate if there was any observable positive relation between the use of high tech gamification and the improved EFL learning performance of Gen Z students. This curiosity had initially stemmed from the seemingly logical presumption that high

tech gamification would be an effective method for the sample groups of this study due to the participants' defining status as students from one of today's most technology-oriented generations, as discussed previously in section 1.2.3 on the relevancy of gamification in the 21st century. Based on the results, there was enough evidence to at the very least suggest that this relation indeed does exist. Although the correlation between high tech gamification and Gen Z students was not directly targeted in this study, the inferred value of the results was that it encourages current and future instructors to be more conscious of how relevant their teaching methods should be in respect to the pupils, who are never characteristically the same as the decade before in terms of natural skills, affinities, and styles of learning. The results of this study implied that high technology is an important factor of today's students' educational context, and that it was through adjusting the teaching methodology to the strengths of the participants that produced stronger learning outcomes. Based on this reasoning, it would be in students' best interest for instructors to try to be more innovative and implement gamified learning activities that incorporate the use of high technology, rather than rely upon traditionally low or no technology based instruction. This may at first be a difficult shift in the system, especially given that most Spanish instructors in K-12 education tend to be from much older generations, but it should not be regarded as a futile attempt. On a positive note, the Bilingual and Cultural Project in Spain has also recruited many native English language assistants into CLIL EFL classrooms across the country; many of these language assistants, such as the researcher herself, happen to be younger millennials who also tend to be strongly familiar and comfortable with the use of many mediums of high technology in their daily lives. Thus, in cases where older generation instructors are open-minded to modernizing their classroom with gamification but struggle or are not as motivated by the use of technology as their young and often technology-obsessed pupils, language assistants could potentially help fulfill these changes, both willingly and more time effectively.

6. Conclusion

6.1. Summary

In this study, the researcher explored gamification in education from two perspectives: low tech gamification and high tech gamification. The first type referred to gamified learning approaches that did not prioritize modern technology as a necessary, structural component of the learning process, while the latter type exclusively did. The importance of making a distinction between these types of gamification was to understand whether technology played a critical aspect of effective gamified learning, especially since most studies experimenting on this topic have commonly assumed that gamification in the 21st century incorporates the use of high technology. Ultimately this investigation of gamification in the classroom was intended as a means to explore and suggest a possible solution towards enhancing EFL classrooms in Spain, a country where the widely implemented foreign language teaching method, CLIL, although well-intentioned, seems to be falling short of adequately addressing the nation's needs and desires to acquire advanced EFL proficiency skills. At the same time, the researcher also hoped that the findings of this study would suggest that there would be an observable relation between high tech gamification and the participants, given that the participants belonged to a new generation in which the use of high technology has become a daily norm. The researcher used a pre and post test method to gather data from two independent samples from the same subject pool, 6th grade Spanish CLIL EFL students; one group was exposed to low tech gamification, and the other, high tech gamification. The researcher then performed *t*-tests on the data, first to check if both forms of gamification increased test scores at all, and secondly to measure if the participants' test scores in the high tech gamified group were statistically significantly higher compared to the low tech group's.

The results of this study generally supported the most recent findings of the use of gamified learning in the educational sector as having an overall positive effect in the classroom. There was evidence that both forms of gamification increased test scores. Moreover, high tech gamification was proven to be statistically significantly more effective than low tech gamification as the researcher had hypothesized. This study has contributed to the literature of gamification in education by providing reason for researchers to establish a stricter classification system of gamification based on the level of technology involved; providing more evidence that gamified learning supports increased learning, especially through high tech gamification; and also by revealing that gamification is a flexible teaching method that can be adapted to both low and

high tech classroom contexts. This is valuable knowledge first and foremost on behalf of students, who directly benefit from the unique and highly engaging learning environments fostered by gamification, and secondly for instructors and educational curriculum designers who are willing to try to modernize their classrooms through gamification, both when there is the infrastructure to cater towards the technological strengths of their students and also when these resources are not available. Finally, this study has contributed to developing understandings of how gamification may function in K-12, foreign language learning contexts, particularly the Spanish CLIL EFL classroom, which has been largely underrepresented in the gamification research community.

6.2. Limitations

One of the main limitations of this study was that the sample represented a very small and narrow group of participants. Therefore, while the results of this study may provide some useful information about this sample of the target population, the results should not be generalized to the greater population of all 6th grade Spanish CLIL/EFL students, nor to other related populations with similar student/learning profiles. Another main limitation was that due to time constraints, the researcher did not provide a qualitative analysis and discussion of the data, which could have offered a more complete and richer picture of the results obtained.

6.3. Further Research

Given the acknowledged lack of a qualitative analysis in this study, future studies of this kind are recommended to implement a mixed methods design. While it is valuable to analyze data from a quantifiable statistical significant perspective, it lacks the full perspective that qualitative analysis can add to investigate any possible underlying reasons influencing the figures at hand. Moreover, further studies in this topic and field could be carried out with larger sample sizes, longer time extended studies, and different subpopulations within the K-12, Spanish CLIL/EFL context, and/or other subject pools entirely. It could also be useful to explore the use of various existing mediums of high and low tech gamification, including testing the effectiveness of other gamified learning websites and applications as well as investigating the kinds of gamified teaching plans and tasks designed by instructors and their impact on the learning process. Future studies could also be primarily designed to test for any correlations between specific factors such as high tech gamification and Generation Z learners, since the

role and use of technology in the classroom will inevitably keep expanding. The exciting reality is that the existing research on gamification in education is so stratified that there are still many gaps that require more time and research, so any new research investigations in this field are highly welcome.

7. Reference List

Alcula. (2019). Statistics calculator: box plot. Retrieved August 30, 2019, from http://www.alcula.com/calculators/statistics/box-plot/

Alcula. (2019). Statistics calculator: Interquartile range. Retrieved August 30, 2019 from http://www.alcula.com/calculators/statistics/interquartile-range/

Cambridge Assessment (2012). Cambridge English vocabulary list, Key English Test (KET), Key English Test for Schools (KETfS) [Educational guide]. Retrieved from https://www.cambridgeenglish.org/Images/22105-ket-vocabulary-list.pdf

Caponetto, I., Ott, M. & Earp, J. (2014). Gamification and education: A literature review. Retrieved from https://www.researchgate.net/publication/266515512_Gamification_and_Education_a_Literature_Review

Baidak, N., Balcon, M.P., & Motiejunaite, A. (2017). Key data on teaching languages at school in Europe [data brief]. Retrieved from https://www.anefore.lu/wp-content/uploads/2017/10/EURYDICE-Key-Data-on-Teaching-Languages-at-School-in-Europe.pdf

Dimock, M. (2019, January 17). Defining generations: Where millennials end and Generation Z begins. Retrieved July 19, 2019, from https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/.

Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International Journal of Educational Technology in Higher Education, 14*(9). DOI: 10.1186/s41239-017-0042-5

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Educational Technology & Society. 18*(3), 75-88. Retrieved from https://www.researchgate.net/publication/270273830_Gamification_in_Education_A_Systematic_Mapping_Study/stats

EF Education First. (2018). *EF EPI. EF English Proficiency Index* (Publication). Retrieved July 23, 2019, from https://www.ef.edu/__/~/media/centralefcom/epi/downloads/full-reports/v8/ef-epi-2018-english.pdf

Eurostat. (2018). Foreign language learning statistics. Retrieved July 29, 2019, from https://ec.europa.eu/eurostat/statistics-explained/index.php/Foreign_language_learning_statistics#Upper_secondary_education

Gamification. (2019). In *Wikipedia.* Retrieved August 10, 2019 from https://en.wikipedia.org/wiki/Gamification

GraphPad. (2018). T-test calculator. Retrieved August 30, 2019, from https://www.graphpad.com/quickcalcs/ttest1/?Format=50

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences*. Retrieved from https://www.academia.edu/4533198/Does_Gamification_Work_A_Literature_Review_of_Empirical_Studies_on_Gamification

Isaacs, S. (2015, January 15) The difference between gamification and game-based learning. Message posted to http://inservice.ascd.org/the-difference-between-gamification-and-game-based-learning/

Kapp, K.M. (2012) *The gamification of learning and instruction.* San Francisco, CA: Pfeiffer.

Kim, S., Song, K., Lockee B., Burton J. (2018). Theories for gamification in learning and education. In Gamification in Learning and Education: Enjoy Learning Like Gaming. *Advances in Game-Based Learning.* (pp. 39-47). Springer, Cham.

Koivisto, J., & Hamari, J. (2019) The rise of motivational information systems: A review of gamification research. *International Journal of Information Management, 45,* 191-210. DOI: 10.106/j.ijin.fomgt.2018.10.013

Landers, R. (2015). Developing a theory of gamified learning. *Simulation & Gaming,* 1-17. DOI: 10.1177/1046878114563660

Lasagabaster, D., & Ruiz de Zarobe, Y. (Eds.) (2010). *CLIL in Spain: Implementation, results and teacher training*. Newcastle: Cambridge Scholars Publishing.

Low-tech.(2019) In *Merriam-Webster Online,* Retrieved July 29, 2019, from https://www.merriam-webster.com/dictionary/low-tech#h1

Majuri, J., Koivisto, J., & Hamari, J. (2018) Gamification of education and learning. *Proceedings of the 2nd International GamiFIN Conference (GamiFIN 2018),* 11-19. Retrieved from https://trepo.tuni.fi/bitstream/handle/10024/104598/gamification_of_education_2018.pdf?sequence=1

Nation, I.S.P. (1983). Testing and teaching vocabulary. *Guidelines, 5*(1), 12-25.

Nicholson, S. (2015). A RECIPE for meaningful gamification. In T. Reiners & L. Wood (Eds.), *Gamification in Education and Business,* 1-20. DOI: 10.1007/978-3-319-10208-5_1

OECD, Organization for Economic Co-operation and Development. (2018). *Teachers by age.* Retrieved July 20, 2019, from https://data.oecd.org/teachers/teachers-by-age.htm

Sailer, M., & Homner, L. (2019) The gamification of learning: A meta-analysis. *Educational Psychology Review.* DOI: 10.1007/s10648-019-09498-w

Seaborn, K., & Fels, D.I. (2015) Gamification in theory and action: A survey. *Int. J. Human-Computer Studies, 74,* 14-31. doi: 10.1016/j.ijhcs.2014.09.006

Shepherd, E., & Ainsworth, V. (2017). *English Impact Una evaluación de la competencias en inglés* (Publication). Retrieved July 23, 2019, from https://www.britishcouncil.es/sites/default/files/british-council-english-impact-report-madrid.pdf

Stangroom, J. (n.d.) Homogeneity of variance calculator – Levene's test. Retrieved August 30, 2019, from https://www.socscistatistics.com/tests/levene/default.aspx

Statistics Kingdom. (n.d.) Shapiro-Wilks calculator. Retrieved August, 30, from http://www.statskingdom.com/320ShapiroWilk.html

Statistics Kingdom. (n.d.) Wilcoxon signed-rank test calculator. Retrieved September 11, 2019, from http://www.statskingdom.com/175wilcoxon_signed_ranks.html

Subhash, S., & Cudney, E.A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior, 87,* 192-206. DOI: 10.1016/j.chb.2018.05.028

Textfixer.com (2019). Random choice generator. Retrieved August 30, 2019 from https://www.textfixer.com/tools/random-choice.php

Zichermann, G. (2011) *How games make kids smarter.* [Video file]. Retrieved from https://www.ted.com/talks/gabe_zichermann_how_games_make_kids_smarter

8. Appendix.

8.1. Pre Test

This is a copy of the pre test used in the study. This copy reflects the 75 questions/vocabulary terms that were officially introduced and studied during the gamified lessons.

KET VOCABULARY PRE TEST

FAMILY

Write F for female, M for male, or B for both.

1. aunt _____
2. cousin _____
3. daughter _____
4. husband _____
5. neighbor _____
6. uncle _____
7. wife _____

FOOD

True or False

1. If you are thirsty, you want to drink water. _____

List the different foods in the correct list.

biscuit bread cheese chicken coffee dessert fish juice omelette

Sweet Foods	Non-sweet Foods	Type of Meat	Drinks
1.	1.	1.	1.
2.	2.	2.	2.
	3.		

ENTERTAINMENT

Match the correct letter with each picture.

a. chess
b. concert
c. instrument
d. museum
e. news
f. magazine
g. newspaper
h. photograph

1.____ 2.____ 3.____ 4.____




5.____ 6.____ 7.____ 8.____




HOUSE/HOME

Match the words to the correct definitions.

1. _____ sitting room	A. a building for cars
2. _____ garden	B. a place to grow plants
3. _____ living room	C. a room where you eat meals
4. _____ dining room	D. bathroom
5. _____ apartment	E. toilet
6. _____ garage	F. a place to live in a shared building
7. _____ bedroom	G. a room in a house where people sit together, watch television, etc.
8. _____ kitchen	H. room to sit and visit with others
9. _____ cupboard	I. a room to sleep
10. _____ bathroom	J. a room for cooking
11. _____toilet	K. a piece of furniture to store food or dishes

PLACES AND BUILDINGS

Match the words to the correct definitions.

1. _____ hotel	A. a place for people to send letters
2. _____ petrol station	B. a place where you pay to sleep on vacation

3. _____ post office	C. a place where people keep their money
4. _____ hospital	D. a place to buy medicine
5. _____ pharmacy	E. a place to buy petrol for your car
6. _____ sports centre	F. a building where you can play different sports
7. _____ bank	G. a place to eat lunch at school
8. _____ cafeteria	H. a place where sick or injured people are given care or treatment

ELECTRONICS/APPLIANCES

True or False

1. A device that is used for playing music CDs → CD player
2. A machine used for keeping food cold → washing machine
3. Portable computer → fridge
4. A machine that washes clothes → washing machine

EDUCATION

Match the words to the correct definitions.

1. ____ classroom	A. someone in your class at school
2. ____ dictionary	B. a place where you can borrow books
3. ____ library	C. a piece of furniture to hold and organize books
4. ____ classmate	D. a reference book that you use to find the definitions
5. ____ bookshelf	E. a classroom object that you can write on with chalk
6. ____ blackboard	F. a room in a school where you have lessons

CLOTHING

Match the words to the correct definitions.

1. _____ belt	a. A short-sleeved casual top
2. _____ dress	b. A band of material that is worn around a person's waist
3. _____ earrings	c. an outer top garment to keep you warm
4. _____ glasses	d. A garment hanging from the waist for girls
5. _____ jacket	e. Clothes for swimming
6. _____ pocket	f. Something you wear over your eyes to improve your vision
7. _____ shoes	g. Glasses that keep your eyes safe from the sun
8. _____ skirt	h. Jewelry for the ears
9. _____ sunglasses	i. One piece of clothing with a top and skirt for a girl
10. _____ swimsuit	j. A small pouch inside a garment for carrying small things

11. _____ T-shirt — k. Something you wear to protect your feet

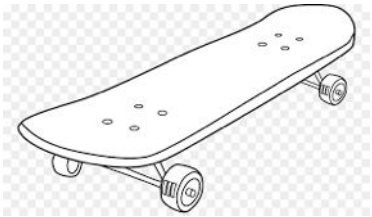
12. _____ umbrella — l. An object that you hold over your head when it is raining

SPORTS

Write the name of the sport under each picture.









8.2. Post Test

This is a copy of the post test used in the study. It consisted of a total of 43 questions based on randomly selected vocabulary terms from the learning content officially covered during the gamified lessons, as shown in the pre test copy.

KET VOCABULARY POST TEST

FAMILY

Write F for female, M for male, or B for both.

1. Uncle _____
2. Husband _____
3. Daughter _____

FOOD

List one correct food in each list.

juice chicken omelette cheese bread dessert

Drinks	Meat	Sweet Food
1.	2.	3.

ENTERTAINMENT

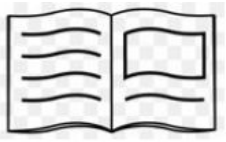
Match the correct letter of the picture to each word.

1. concert _____
2. news _____
3. chess _____
4. photograph _____

A. B. C.

D. E. F.





HOUSE/HOME

Match the words to the correct definitions.

1. _____ garage	A. a place to grow plants
2. _____ cupboard	B. a room to sleep
3. _____ garden	C. a room for cooking
4. _____ apartment	D. a place to live in a shared building
5. _____ toilet	E. a place where you eat meals
6. _____ sitting room	F. a room where you sit to talk to your visitors
	G. a room where you sit to watch TV
	H. a bathroom
	I. a building to park a car
	J. a piece of furniture for storing food or dishes

PLACES AND BUILDINGS

Match the words to the correct definitions.

1. _____ petrol station	A. a place to keep your money
2. _____ cafeteria	B. a place to send letters and postcards
3. _____ sports centre	C. a place to go when you are sick or injured
4. _____ bank	D. a place to play different sports
5. _____ hotel	E. a place to buy medicine
6. _____ post office	F. a place to sleep when you are on vacation
	G. a place to eat lunch at school
	H. a place to fill up your car with gas

ELECTRONICS/APPLIANCES

True or False.

1. A device that is used for playing movies and films → CD player
2. A portable computer → laptop
3. A machine that washes clothes in cold water → fridge

EDUCATION

Match the words to the correct definitions.

1. ____ blackboard	A. someone in your class at school
2. ____ library	B. a place where you can borrow books
3. ____ classroom	C. a piece of furniture to hold and organize books
4. ____ classmate	D. a reference book that you use to find the definitions
5. ____ bookshelf	E. a classroom object that you can write on with chalk
	F. a room in a school where you have lessons

CLOTHING

Write the correct word to match the definitions.

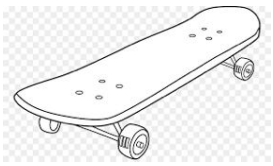
shoes earrings jacket belt dress skirt

sunglasses T-shirt swimsuit umbrella pocket glasses

1. You wear these over your eyes to help you see better →
2. You wear these to protect your feet →
3. You wear these to protect your eyes from the sun →
4. You wear this when you swim →
5. A garment for girls that hangs down from the waist →
6. A small pouch in your clothes used to carry small things →
7. An outer top garment to keep you warm →

SPORTS

Write the name of the sport under the picture.



8.3. Final Vocabulary List

This was the final list of 43 vocabulary terms assessed and marked in both the pre and post tests. Only these terms were considered applicable during the final data collection, analysis procedures, results, and discussions; they are organized by topic category:

Family	Food	Sports
Uncle	Chicken	Badminton
Daughter	Dessert	Basketball
Husband	Juice	Skateboarding
		Skiing
Entertainment	House/home	Swimming
Chess	Apartment	Tennis
Concert	Cupboard	
News	Garage	
Photograph	Garden	
	Sitting room	
Places/Buildings	Toilet	
Bank		
Cafeteria	Electronics/appliances	
Hotel	CD player	
Petrol station	Laptop	
Post office	Washing machine	
Sports centre		
	Clothing	
Education	Dress	
Blackboard	Glasses	
Bookshelf	Pocket	
Classmate	Shoes	
Classroom	Skirt	
Library	Sunglasses	
	Swimsuit	