EXAMINING EFL STUDENTS' MOTIVATION AND ATTITUDES TOWARD A GAMIFIED COURSE USING LEADERBOARDS AND QUESTS AT A JAPANESE UNIVERSITY

A Thesis submitted by

Andrew Philpott

For the award of

Doctor of Philosophy

2020

Abstract

This thesis reports on an investigation into the use of leaderboards and quests (i.e., quest-based learning) in a gamified English as a foreign language (EFL) course at a Japanese university. The study focuses on gamified instructional design, a type of gamification that incorporates specific components of traditional games into the structure of an academic course to influence student behaviour. The study uses Deci and Ryan's (1985) Self-determination theory (SDT) as the theoretical foundation to explore how leaderboards and quests affect student performance (i.e., amount of work completed) and foreign language (FL) motivation. For the study, the researcher developed a gamified EFL course using leaderboards and quests as central gamification components. The main part of the study was conducted over a 14-week period with two intact classes of participants: Class 1 (n = 26) and Class 2 (n = 20). The leaderboard acted as an independent variable as it was only used in Class 1. A quasi-experimental mixed methods research design was utilised to collect and analyse data from five data collection instruments (i.e., performance-related data, a leaderboard questionnaire, quest diaries, the Language Learning Orientations Scale (LLOS), and semi-structured interviews) to answer three research questions.

The first research question examined leaderboards to determine how they affect student performance. The results of the study showed that the participants generally enjoyed the leaderboards. Four ways the leaderboards affected performance were identified: (1) they encourage performance but limit performance; (2) they could negatively impact learning outcomes if they reward quantity over quality; (3) they positively affect the emotions, attitudes, and performance of the participants who have a middle or high leaderboard rank; (4) they negatively affect the emotions, attitudes, and performance of students who have a low leaderboard rank. The second research question explored the participants' opinions and perceptions of quest-based learning (QBL) to determine its viability as a pedagogical approach for EFL courses. The results showed overwhelmingly positive opinions and perceptions towards QBL, and provided evidence that QBL can increase intrinsic FL motivation while supporting FL learning. The third research question examined the effect leaderboards and QBL had on the participants' FL motivation. The results showed that

leaderboards increase extrinsic FL motivation by using points and rank to control behaviour, and undermine intrinsic FL motivation more than they support it. The quests increased the participants' intrinsic FL motivation. The increase was attributed to the quest design that leveraged specific aspects of multiple motivation theories (e.g., SDT, positive psychology, international posture).

The findings from this thesis make several important contributions to the bodies of literature surrounding gamification. First, the findings provide an updated perspective on the state of FL motivation at a Japanese university. Second, the findings show that SDT is an effective theory to analyse gamified instructional design implementations. SDT was able to align leaderboards to external regulation and show that leaderboards shift internally leaning extrinsic motivation to externally grounded extrinsic motivation. Third, based on all the findings, the thesis introduces a new framework to guide future gamified instructional design implementations. Fourth, the thesis details how QBL can be used for EFL pedagogy to support FL learning and intrinsic FL motivation. The thesis concludes with suggestions for future research that aims to replicate the results of this study, and research that examines how different types of leaderboard configurations affect performance and FL motivation.

Certification of Thesis

This Thesis is entirely the work of Andrew Philpott except where otherwise acknowledged. The work is original and has not previously been submitted for any other award, except where acknowledged.

Principal Supervisor: Associate Professor Jeong-Bae Son

Associate Supervisor: Associate Professor Ann Dashwood

Student and supervisors signatures of endorsement are held at the University.

Acknowledgements

Writing this thesis has been a long and challenging experience. Analysing my own motivation to complete a PhD provided informative insight into this motivation-related study. Completing this thesis required more than my own personal motivation, and I want to express my appreciation to a number of people who have helped me through this journey.

First, I would like to thank my primary thesis supervisor Associate Professor Jeong-Bae Son whose expertise, guidance, and encouragement was invaluable throughout all stages of this thesis. Second, I would like to thank my secondary supervisor Associate Professor Ann Dashwood for providing timely feedback. I would like to thank the students who were the participants of this study.

Next, I would like to thank those who helped me with specific parts of the thesis: Mimi Kihara for emotional support and translating, Ran Niboshi for conducting the interviews, Dai Abukawa for transcribing the interviews, Simon Bibby for introducing me to Lee Sheldon's book "The Multiplayer Classroom", and Scott Aubrey for his friendly advice and encouragement throughout the thesis. I would also like to thank Judd Martin, Ryan Pain, Kristoffer Huupponen, Jon Watkins, David Lees, Andrew Nowlan, Abdullah Alami, Craig Gamble, and Andrew Went for their insightful feedback and support.

Finally, I would like to thank my parents Robert and Iris Philpott for the wide range of opportunities, love, and support they have provided for me.

This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

Table of Contents

List of Tables

List of Figures

Abbreviations

The following abbreviations have been used. The abbreviations have been explained in Section 1.3.

AL	Autonomous learning
CALL	Computer-assisted language learning
CET	Cognitive Evaluation Theory
EC	English Central
EFL	English as a foreign language
ER	Extensive reading
ESL	English as a second language
FL	Foreign language
GID	Gamified instructional design
IP	International posture
L2	Second language
L2/FL	Second and/or foreign language
L2MSS	L2 Motivation Self System
MDA	The Mechanics-Dynamics-Aesthetics framework
OIT	Organismic Integration Theory
PP	Positive psychology
QBL	Quest-based learning
SDT	Self-Determination Theory
SLA	Second language acquisition
TBLT	Task-based language teaching
TL	Target language
TOEFL	Test of English as a Foreign Language
TOEIC	Test of English for International Communication
XP	Experience points

Chapter 1. Introduction

1.1 Background

The motivation for this study stemmed from the researcher's need for an innovative approach to encourage students to complete all their homework in an English as a foreign language course (EFL) at a Japanese university. Japanese students' lack of motivation towards studying English is well-known and often researched in the field of applied linguistics (e.g., McVeigh, 2004; Nakata, 2006); this is because motivation is considered a crucial factor in determining a student's foreign language success (Gardner, 1985; Gardner & Lambert, 1959). Researchers continually explore new theories, models, and innovative approaches to engage students in the process of learning English (e.g., Dörnyei & Ushioda, 2011). Based on research findings, dedicated English language teachers employ suitable pedagogical approaches and activities such as task-based language teaching (TBLT), autonomous learning, or computer-assisted language learning (CALL) to motivate their students. Gamification is a new approach that can be applied to language courses to potentially motivate students to not only engage with course content but also foster their long-term intrinsic motivation towards improving the English ability (Sheldon, 2011).

Gamification is commonly mistaken for game-based learning or game theory (Cózar-Gutiérrez & Sáez-López, 2016; Eng, 2019). Game-based learning, more specifically, digital game-based learning, is the concept of playing video games to learn or practice something that has an educational objective (e.g., Gee, 2003; Kim, Park, & Baek, 2009; Prensky, 2001; Reinhardt 2017, 2019); for example, *Minecraft* to teach students how to build structures (Farber, 2016), or *World of Warcraft* as an avenue to practice a foreign language (Soares Palmer, 2010). Game theory is "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers" (Myerson, 1999, p. 1); The Prisoner's Dilemma is a famous example of game theory that shows how an optimal solution occurs when decision-makers cooperate with each other. Gamification, on the other hand, refers to the use of game elements and game design techniques in non-game contexts to motivate desired behaviours (Deterding, Sicart, Nacke, O'Hara, & Dixon, 2011; Werbach & Hunter, 2012). After being first mentioned in a blog post in 2008 (Terrill, 2008),

implementations of gamification in a variety of fields increased quickly due to new technologies being able to track, analyse, and display data (Deterding, 2012).

Gamification is used to bridge the gap between current behaviour and desired behaviour (Werbach & Hunter, 2012). It can be used in positive ways to encourage people to live more safely, healthily, or socially (e.g., McGonigal, 2011; Xu, 2012), or used in negative ways to influence people to work harder than necessary or increase their product purchasing behaviour (e.g., Harwood & Garry, 2015; Werbach & Hunter, 2012). The Nike+ Running App in 2010 was one of the first and most famous examples of successful gamification (Deterding, 2012). The Nike+ App motivated people to run more by using gamification elements such as leaderboards, achievements, and social media to track, analyse, and compare individuals' performance data with their previous results and other users of the system. The Nike+ gamification implementation was successful because it encouraged people to live a healthy lifestyle, leveraged smartphone and social media obsession, and allowed Nike to project a positive company image while increasing their profits (Codish & Ravid, 2014; Werbach & Hunter, 2012; Zichermann & Cunningham, 2011).

In contrast to Nike's implementation, a Disneyland hotel in California was criticised for exploiting their employees using gamification. Their system used public monitors to display and compare the work speed of housekeeping employees. The system compared the speed of employees loading pillowcases, sheets, and other items into a laundry machine. If the expected work rate was near or above 100%, their names were displayed in green, if not, their names were displayed in red. This system manipulated employees to work harder than necessary and led to negative feelings about the workplace becoming too competitive with some employees skipping bathroom breaks to ensure their efficiency rating would not decline. Employees named this system the electronic whip (Lopez, 2011).

Proponents of gamification (e.g., Chou, 2015; González & Area, 2013; McGonigal, 2011; Muntean, 2011; Tan & Hew, 2016; Werbach & Hunter, 2012) argue that gamifying an activity makes the activity more engaging and motivating. In contrast, gamification critics consider gamification as "the high fructose corn syrup of

engagement" (Sierra, 2011, para. 14), *exploitationware* (Bogost, 2011), *pointsification* (Robertson, 2010), a *hedonic treadmill* (Werbach & Hunter, 2012), and a *novelty* that cannot sustain engagement (Cook, 2018). Bogost (2011) asserts that gamification replaces real, meaningful incentives with fictional, superficial incentives; psychologists (e.g., Frey & Jegen, 2001) refer to this in motivation crowding theory as 'crowding-out', a situation in which the extrinsic rewards crowd out the intrinsic rewards, with no guarantee that the intrinsic reward will be remaining once the extrinsic reward has been removed or devalued. On the surface, it appears that the critics may be correct, as many gamification implementations only use extrinsic rewards such as points, badges, and leaderboards to drive behaviour; however, the proponents also warn that systems that rely on extrinsic rewards will not sustain engagement or foster meaningful engagement. The proponents suggest leveraging relevant aspects of psychological theories such as behaviourism, self-determination theory (SDT), and positive psychology to design effective gamification systems that can achieve resilient behaviour change and avoid situations in which short-term spikes in performance lead to future demotivation (Hamari, Koivisto, & Sarsa, 2014).

Gamified instructional design (GID) is a type of gamification that refers to the concept of applying game components and mechanics to the structure or content of an educational course to motivate students to engage with the learning content. Gamifying an academic course does not require digital technology; however, GID can benefit from the digital ability to track, share, and analyse data (Buckley & Doyle, 2014). Sheldon (2011) explains that GID requires the teacher to redesign various elements of the course to be more game-like; for example, assignments are presented as quests, and when a student completes a quest, they are rewarded with points that are displayed on a class leaderboard. In this case, an underlying assumption could be that, because leaderboards foster competition, when the students see their leaderboard ranking, they will feel motivated to perform better.

There are over 100 different gamification techniques that can be uniquely calibrated, independently applied, and interconnected with other gamification techniques to attempt to achieve the specific goals of an implementation (Dubravac, 2012). The amount and visibility of the gamification applied to a gamified course can vary from

very little, in which students are unaware that something has been gamified, to a lot, in which the course appears very game-like. Richter, Raban, and Rafaeli (2015) explain that it is not clear how certain combinations of gamification techniques will affect student performance and motivation. Therefore, if the assertion that gamification increases performance and motivation (Chou, 2015; Sheldon, 2011; Werbach & Hunter, 2012) is true, research examining different combinations of gamification techniques is required to guide future GID implementations.

Empirical evidence supporting gamification is mixed (Koivisto & Hamari, 2019); for the educational context, the literature is lacking (Hanus & Fox, 2015; Reinhardt, 2019; Tan & Hew, 2016; Winoto & Yang, 2015) and sometimes conflicting (Domínguez et al., 2013; Garland, 2015; Lee & Hammer, 2011). An underlying theme that emerges from the literature (e.g., Barata, Gama, Jorge, & Gonçalves, 2017; Buckley, Doyle, & Doyle, 2017; Garland, 2015; Hamari et al., 2014) is the importance of context; different students respond differently to the same gamification components, and what was effective in one situation might not be effective in a different situation. This uncertainty about the use of gamification in educational contexts led to many researchers calling for empirical studies that:

1. Isolate gamification components such as leaderboards or quests (e.g., Conley & Donaldson, 2015; Devers & Gurung, 2015; Diamond, Tondello, Marczewski, Nacke, & Tscheligi, 2015; Landers, Bauer, Callan, & Armstrong, 2015; Richter et al., 2015);

2. Explore how gamification affects short-term performance and long-term motivation (e.g., Attali & Arieli-Attali, 2015; Barata et al., 2017; Codish & Ravid, 2017; Garland, 2015; Ibanez, Di-Serio, & Delgado-Kloos, 2014; Glover, 2013; Winoto & Tang, 2015);

3. Prove how effective gamification is in specific contexts such as language courses (e.g., Devers & Gurung, 2015; Hamari et al., 2014; Kapp, 2015); and

4. Foster theory development (e.g., Hamari et al., 2014; Kapp, 2015; Thiebes, Lins, & Basten, 2014).

Two gamification components that are of interest to gamified EFL courses are leaderboards and quests. Leaderboards are one of the most common methods to gamify a course (Dicheva, Dichev, Agre, & Angelova, 2015). Literature (e.g., Aldemir, Celik, & Kaplan, 2018; Barata et al., 2017; Poondej & Lerdpornkulrat, 2016) shows that leaderboards have successfully increased the amount of work students complete. This could be beneficial for language courses as second language acquisition (SLA) theory (e.g., Krashen, 2009) suggests that the more students perform in the target language (TL), the quicker L2 acquisition occurs. However, conflicting literature (e.g., Frey & Jegen, 2011; Deci, Koestner, & Ryan, 1999) warns that the short-term performance gains from an extrinsically rewarding leaderboard may come at the cost of students L2/FL intrinsic motivation.

In contrast to the extrinsically focused leaderboards, quests, also referred to as quest-based learning (QBL), show potential to be pedagogically-effective learning activities that foster intrinsic motivation by incorporating psychological theories such as the SDT (Deci & Ryan, 1985) into their design. Leaderboards and quests appear to be diametrically opposed; leaderboards are extrinsic in nature; quests are intrinsic in nature. The combination of leaderboards and quests could potentially allow leaderboards to increase short-term performance while quests foster intrinsic FL motivation, counteracting any negative affects the extrinsically rewarding leaderboards have on intrinsic motivation. Further research examining the combination of leaderboards and quests in a gamified EFL course is required.

1.2 Research Aims and Questions

The study reported in this thesis aims to determine how two gamification components, leaderboards and quests, affect students' performance and FL motivation in an EFL course at a Japanese university. To achieve the aims, the following research questions are addressed:

1: What effect do leaderboards have on student performance in an EFL course?

2: What are students' opinions and perceptions of quest-based learning in an EFL course?

3: What are the effects of leaderboards and quest-based learning on FL motivation?

Research Question 1 (RQ1) aims to determine if and how leaderboards affect student performance in a gamified EFL course. Performance refers to the amount of work students complete. Many gamification studies (e.g., Landers et al., 2015; Tan & Hew, 2016) show that leaderboards positively impact student performance; however, the studies fail to isolate the specific aspects of the leaderboards that affect performance. Research Question 2 (RQ2) explores students' opinions and perceptions of QBL in a gamified EFL course. Students' opinions refer to what students think about specific aspects of QBL, whereas their perceptions refer to how they regard QBL as an approach to learning EFL. RQ2 aims to determine the viability of QBL as an intrinsically motivating approach to EFL pedagogy. Research Question 3 (RQ3) investigates the effect leaderboards and quests have on FL motivation. Special attention is paid to see if the concern (e.g., Bielik, 2012) that leaderboards negatively affect intrinsic motivation, and the suggestion (e.g., Sheldon, 2011) that quests positively impact intrinsic motivation, is reflected in the data.

1.3 Terms and Definitions

This thesis uses various acronyms from the applied linguistics research field to denote commonly known concepts. The term target language (TL) refers to the language someone is trying to learn. Second language acquisition (SLA) research examines the factors associated with learning a second language (L2) in the country the TL is natively used; if the TL is English, the term English as a second language (ESL) is used. SLA findings are often extrapolated to learning a foreign language (FL) in a country in which the TL is non-native; in this case, if the TL is English, the term English as a foreign language (EFL) is used. Therefore, this study occurs in an EFL environment as the participants are Japanese students learning English in Japan. Due to the intertwined nature of second/foreign language research, the thesis often

uses the L2/FL acronym to refer to both concepts at the same time; for example, L2/FL motivation refers to motivation to learn a second or foreign language.

The thesis often refers to L2/FL pedagogical approaches and activites to support the design of the gamified course. For example, the communicative language teaching approach of task-based language teaching (TBLT), also known as task-based language learning, is the concept of using tasks to teach and learn a foreign language (Ellis, 2003). Computer-assisted language learning (CALL) is the field of academic enquiry that examines the use of computers and other digital devices in language learning and teaching (Son, 2018). Extensive reading (ER) is a pedagogical approach that aims to improve L2/FL ability through reading in the TL (Day & Bamford, 1998; Krashen, 2009). Graded readers are books designed for this purpose. Autonomous learning (AL) literature (e.g., Benson, 2011) supports the design of the gamified course. Dörnyei's (2005) L2 Motivational Self System (L2MSS) framework, Yashima's (2002) international posture construct (IP), and positive psychology (PP) literature (e.g., Seligman, 2013) provide guidance for designing quests that are intrinsically motivating.

English Central (EC) (http://www.englishcentral.com) is a website designed for students to improve their English language ability while watching videos and doing corresponding activities. The activities aim to foster vocabulary acquisition and improve students' pronunciation and listening skills. The videos generally come from everyday life situations in English speaking countries and range in genre from news reports and famous speeches, to movie trailers and commercials, to TOEIC/TOEFL preparation materials. There are over 10000 different videos which range in difficulty from beginner to advanced. The website can suggest appropriate level content for the student to study or students can choose videos they are interested in. The website can be accessed through a web browser or smart phone app. Teachers can set goals for their students to achieve.

Deci and Ryan's (1985) Self-Determination Theory (SDT) is commonly referenced throughout the thesis to explain L2/FL motivation and the psychology surrounding gamification. SDT is based on Cognitive Evaluation Theory (CET) and Organismic

Integration Theory (OIT). From a L2/FL perspective, CET's main premise is that learning activities that foster the psychological needs of autonomy, competence, and relatedness lead to an increase in the students' L2/FL motivation. Intrinsic motivation refers to wanting to perform some type of behaviour "for its own sake in order to experience pleasure and satisfaction, such as the joys of doing a particular activity or satisfying one's curiosity" (Dörnyei & Ushioda, 2011, p. 23). The thesis often refers to the three psychological needs of autonomy, competence, and relatedness as the tenets of SDT.

Organismic integration theory outlines the three different types of motivation someone can possess: amotivation, extrinsic motivation, and intrinsic motivation. OIT focuses on the different types of extrinsic motivation based on their perceived locus of causality. Extrinsic motivation refers to "performing a behaviour as a means to some separable end, such as receiving an extrinsic reward (e.g., good grades) or avoiding punishment" (Dörnyei & Ushioda, 2011, p. 23). Amotivation refers to a situation in which a student has no motivation to perform an activity because there is "no relation between their actions and the consequences of those actions; the consequences are seen as arising as a result of factors beyond their control" (Noels, Pelletier, Clément, & Vallerand, 2000, p. 62). The thesis uses a questionnaire entitled the Language Learning Orientations Scale (LLOS) to measure students' L2/FL motivational orientation from a SDT perspective.

Identifying, categorising, and understanding gamification is currently a slightly confusing endeavour (Chou, 2015; Werbach & Hunter, 2012). As the concept of gamification is still relatively new, various definitions and descriptions are found in the literature. Based on Deterding et al.'s (2011) and Werbach and Hunter's (2012) definitions, the researcher defines gamification as *the use of game elements and game design techniques in non-game contexts to motivate desired behaviours*. In the study, the non-game context is an EFL course at a Japanese university; the desired behaviour is to motivate students to complete all their homework and to increase their intrinsic FL motivation; and the game design elements and techniques of competition, cooperation, leaderboards, points, and quests are used to try and achieve the goals. Gamified instructional design (GID) is the term given to academic courses that have had gamification applied to their structure or content. Students in a

gamified class could be referred to as players, participants, or students. Points, sometimes referred to as experience points (XP), are used in a gamified environment to reward the player. Quest-based learning (QBL) is the pedagogical approach of designing and delivering learning content as game-like quests (Haskell, 2013).

1.4 Structure of Thesis

Chapter 1 of this thesis introduces the background of the study reported in the thesis, the research aims and questions, the terms and definitions, and outlines the structure of the thesis. Chapter 1 explains how gamification has been successfully applied to a variety of contexts to motivate behaviour; however, literature supporting its application to an EFL course in the Japanese context is lacking and required. Chapter 2 provides a thorough analysis of the literature related to the study and situates the study in the L2/FL motivation field of applied linguistics research. SDT is discussed in detail as the study uses it as the central theory to analyse how gamification effects performance and motivation. Discussion then turns to the catalysts of the recent gamification boom and an argument is presented that EFL courses are suitably positioned to be gamified due to recent advancements in CALL. Chapter 2 then discusses how the gamification literature suggests that leaderboards and quests have potential to increase performance and intrinsic FL motivation for Japanese EFL students; counter arguments are also presented. Finally, the theoretical framework shows how the gamification components are combined with motivation and SLA-related theories to create a gamified course.

Chapter 3 details the research methodology of the study. The first sections explain the quasi-experiential mixed methods research design, the Japanese EFL participants, and the important aspects of how the gamified course was designed. The next section explains the five data collection instruments that were employed to collect data about different aspects of the gamified course. The following sections detail the corresponding data collection and data analyses procedures that occur for each data set. Chapter 3 finishes by covering the ethical considerations of the study, such as the possible emotional risk to the participants, data privacy, and the informed consent procedure.

Chapter 4 presents the results of the data analyses. Section 4.2 presents the results of the quantiative data analyses that were performed the performance-related data. Section 4.3 presents the results of the various data analyses that were conducted on the leaderboard questionnaire data. Sections 4.4 and 4.5 present the results of the quantitative content analyses that were conducted on the quest diaries and the semi-structured interviews. Section 4.6 presents the results of the quantitative data analyses that were performed on the LLOS data. Chapter 5 discusses the results in relation to the research questions and the existing literature. Chapter 6 concludes the thesis with a summary of the important findings, and a discussion of the implications, limitations, and possible directions for future research.

Chapter 2. Literature Review

2.1 Overview

This chapter explores the literature surrounding the use of gamification in the EFL context. Section 2.2 situates this study in the L2/FL motivation research field by examining the major findings that have occurred over the last 60 years. Sections 2.2.1 and 2.2.2 introduce the self-determination theory (SDT) and the Language Learning Orientations Scale (LLOS) as they provide the theoretical foundation to analyse various aspects of a gamified EFL course. Section 2.2.3 details L2/FL motivation in the Japanese context from a historical and contemporary perspective to provide a rationale for this gamification study to occur.

Section 2.3 provides a background as to what gamification is, where it has come from, and how it has been used in a variety of contexts to motivate behaviour. Section 2.3.1 describes how Werbach and Hunter's (2012) framework informs the selection and calibration of gamification components to achieve the specific goals of a gamification implementation. Section 2.3.2 presents literature from the computer-assisted language learning (CALL) and autonomous learning research fields to argue that FL courses are uniquely suitable to be gamified. Section 2.3.3 provides a general introduction to the gamified instructional design research that has occurred so far and provides a rationale for further research that focuses on leaderboards and QBL.

Section 2.4 discusses the use of leaderboards in gamified instructional design with a focus on how they affect student performance and FL motivation. Behaviourism and SDT provide theoretical insight into how leaderboards use rewards such as points and leaderboard rank to influence behaviour. Section 2.4.1 summarises the literature about the impact of leaderboards on student performance and argues that leaderboards can increase student performance in a range of ways; however, performance increases may be unsustainable or limiting, have a negative impact on learning, and only occur for high performing students. Section 2.4.2 summarises the literature about the impact of leaderboards on FL motivation and suggests that if students perceive a leaderboard to be an instrument of control, it will increase their FL extrinsic motivation and decrease their FL intrinsic motivation; however, if

students perceive a leaderboard to be an instrument that provides meaningful information, it can positively affect intrinsic FL motivation. Section 2.4.3 expands the scope of the literature review by asserting that an analysis of students' emotional reactions towards leaderboards could provide insight into how leaderboards affect performance and motivation.

Section 2.5 discusses the potential of QBL to be an intrinsically motivating approach to EFL pedagogy. Due to the dearth of literature that explores quests in an EFL context, Section 2.5.1 explains how SLA theory and task-based language teaching (TBLT) literature can inform the design of quests that support L2/FL acquisition. Section 2.5.2 reviews the literature about the use of quests in non-EFL contexts and suggests that quests can be intrinsically motivating if they incorporate SDT and positive psychology. Section 2.5.3 provides a rational for research to occur that analyses students' opinions and perceptions towards QBL.

Section 2.6 presents the theoretical framework used in this thesis. The theoretical framework is based on the relevant findings that emerged from this literature review. The framework shows how gamification literature, psychological theories, and SLA theories and approaches are merged to support the design of the gamified course that is used in this study. The framework shows how this study collects mixed methods data to determine the effect of leaderboards on performance and FL motivation, and the viability of QBL as an intrinsically motivating and pedagogically effective approach to teaching EFL.

2.2 Motivation in Second/Foreign Language Learning

Gamification aims to motivate behaviour. The word *motivation* derived from the Latin verb *motivus* which means a *moving cause*; this suggests that there are underlining reasons that explain someone's motivation (Cofer & Petri, 2001). The amount of literature that explores the relationship between gamification and motivation in an EFL context is lacking; however, literature from the well-established mainstream psychology and L2 motivation research fields provide guidance to understanding education-based gamification.

From a research perspective, motivation is concerned with the direction and magnitude of human behaviour (Dörnyei & Ushioda, 2011). Early research from notable psychologist Sigmund Freud (1966) examined motivation as an emotional and natural human instinct deriving from the unconscious mind. Later research focused on conscious cognitive processes such as goal-setting, expectations, and self-efficacy beliefs in affecting behaviour (Dörnyei & Ushioda, 2011). Motivation has been widely studied in mainstream psychology since the 1930s (Allport, 1937; Murray, 1938) to understand what drives a person to engage or not engage in an activity, make certain choices while doing an activity, and why they persist or stop doing an activity.

Many definitions, theories, and frameworks have been developed to conceptualise and detail the concept of motivation in mainstream psychology. However, as the word motivation is so rich in meaning and complex in nature, no one theory can comprehensibly cover all facets of it; varying degrees of suitability to a specific context are present in all theories (Dörnyei & Ushioda, 2011). Motivation theories can be divided into three categories: (1) hedonic or pleasure; (2) cognitive or need-to-know; (3) growth or actualization (Roeckelein, 2006). Hedonic or pleasure motivation theories such as Herzberg's motivation theory (Herzberg, Mausner, & Sydnderman, 1959), attribution theory (Heider, 1958), and opponent-process theory (1974) posit that human behaviour is motivated by motivating stimuli. Cognitive or need-to-know motivation theories such as cognitive dissonance theory (Festinger, 1957), expectancy theory of motivation (Vroom, 1964), and goal-setting theory of motivation (Locke & Latham 1990) "posit that motivation is the result of active information-processing where an individual, subconsciously or consciously positively evaluates the acting out of a specific behavior, thus is motivated" (Leadership-central.com, 2018, para. 4). Growth or actualization motivation theories such as Maslow's (1943) hierarchy of needs, Alderfer's (1969) ERG theory, and Deci and Ryan's (1985) SDT explain that humans are motivated when involved in activities that lead to personal growth, self-fulfilment, and self-actualisation.

Motivation is considered a crucial determinant of learning and it has been widely explored in the general education research field (Buckley & Doyle, 2014). The

debate in the education literature as to whether motivation is a cause, or an effect of learning has moved to an agreement that motivation is cyclic in nature with learning; high motivation leads to high achievement which in turn leads to high motivation; the same cycle is present for the negative alternative (Dörnyei & Ushioda, 2011). Researchers (e.g., Dweck, 1999; McCombs, 1994) have focused on ways to break the negative cycle by "modifying the cognitive processes (e.g., learners' self-perceptions and interpretations of events) that mediate the relationship between motivation and learning" (Dörnyei & Ushioda, 2011, p. 6). However, accurately measuring motivation is difficult as an individual's cognition, behaviour, and achievement are affected by their physical, cultural, and psychological context (Dörnyei & Ushioda, 2011).

The goal of L2/FL language teachers is to increase students' L2/FL proficiency. Applied linguistic researchers (e.g., Apple, Da Silva, & Fellner, 2013; Bernaus & Gardner) examine the factors that can improve proficiency. Motivation is one of the most studied elements of SLA as it is considered a crucial factor in determining second/foreign language success (Brown, 2000; Buckley & Doyle, 2014). Dörnyei and Ushioda (2011) explain that L2/FL motivation research borrows from the wealth of knowledge already established in the mainstream psychology research field; the knowledge is adapted to account for the unique social and psychological characteristics of L2/FL learners. SLA research before 1959 supported the belief that intelligence and aptitude were the main determinants of successful L2 acquisition (e.g., Carroll, 1958). Since 1959, L2/FL research evolved through three district phases: the social psychological period (1959–1990), the cognitive-situated period (during the 1990s), and the process-oriented period (from 2000) (Dörnyei, 2005).

The social psychological period includes numerous academic theories and frameworks that attempt to identify and explain the factors affecting L2 performance. This period focused on ethnolinguistic minority communities learning the language of a dominant community; with findings later extrapolated to the EFL research field. Clément, Gardner, and Smythe (1977) theorised the importance of linguistic self-confidence for learning and using the language of a dominant community. Giles and Byrne's (1982) intergroup model provided a framework to understand how factors such as perceptions and identification affect ethnic minority groups learning a L2.

Schumman's (1978) acculturation theory focused on the process of individual acculturation into a dominant community and detailed how social and psychological distance between the language learner and the target language group is detrimental to L2 language achievement. However, it was the work of Gardner and Lambert (1972) that is synonymous with the social psychological period of L2 motivation research.

The social psychological research period grew from Lambert's (1955, 1956a, 1956b, 1956c) studies in the second half of the 1950s; however, Dörnyei and Ushioda (2011) explain that it is Gardner and Lambert's (1959) study of high school students in Montreal learning French as an L2 that signals the beginning of the social psychological period. This initial study was the first investigation into individual differences in relation to the impact of L2 learners' attitudes and motivation on L2 achievement. Gardner and Lambert (1959) found two factors related to language achievement: (1) linguistic aptitude, and (2) motivation. They characterised motivation "as a willingness to be like valued members of the language community" (Gardner & Lambert, 1959, p. 271). This work led to their seminal publication (Gardner & Lambert, 1972) which showed a learner's attitude toward the L2 community significantly affects their L2 learning behaviour. In contrast to the cognitive theories being explored in mainstream psychology during this time, Gardner and Lambert's (1959, 1972) studies showed that language learning is a social process with unique characteristics that are not present in the learning of other subjects.

Based on Gardner and Lambert's (1972) findings, Gardner and Smyth (1975) developed the socio-educational model of second language acquisition; a model that has been revised many times (e.g., Gardner, 1979; Gardner, 1985; Gardner & McIntyre, 1993; Tremblay & Gardner, 1995; Gardner 2010). Each iteration of the model led to more complexity, the incorporation of external factors, and more characteristics explaining each construct. The model (Gardner, 2010) comprises six constructs: language aptitude, attitudes toward the learning situation, integrativeness, motivation, language anxiety, and language achievement. The model outlines how language achievement is determined based on one's ability and motivation, with language learning anxiety playing a role. At the heart of the model is the integrativeness construct; initially referred to as *integrative motive*, it is defined as

"motivation to learn a second language because of positive feelings towards the community that speaks the language" (Gardner, 1985, pp. 82–83).

To accompany the model, Gardner (1985) developed the Attitude/Motivation Test Battery to quantitatively measure four factors involved in L2/FL learning: integrativeness, attitudes toward learning situation, motivation, and language anxiety. It has been widely used (e.g., Bernaus & Gardner, 2008; Gardner & MacIntyre, 1991; Inbar, Donitsa-Schmidt & Shohamy, 2001) as a reliable psychometric tool to explore motivation to learn a foreign language. Research (e.g., Dörnyei & Clément, 2001) has shown that integrativeness is a stronger form of motivation than instrumental orientation. Dörnyei and Ushioda (2011) explain that "instrumental orientation is the utilitarian counterpart of integrative orientation in Gardner's theory, pertaining to the potential pragmatic gains of L2 proficiency, such as getting a better job or a higher salary." (p. 41). This has resulted in the concept of integrativeness becoming highly influential and often researched in the SLA research field.

Even though the socio-educational model was the first and most influential theory of L2 motivation, criticism towards it (e.g., Oller, Hudson, & Liu, 1977) started soon after the seminal 1972 paper (Gardner & Lambert, 1972) was published. Lukmani (1972) found that Indian students with a strong instrumental orientation achieved higher levels of language proficiency compared to students with a strong integrative motivation. Dörnyei (1990) and Horwitz (1990) supported the finding for EFL learners in Hungary and the Philippines respectively. Au (1988) doubted that integrativeness was an antecedent of motivation and suggested that it is only a by-product of L2 achievement.

The focus on integrativeness in the socio-educational model leads to a lack of accounting for a range of cognitive and social factors. For example, several studies (e.g., Clement & Kruidenier, 1983; Ely, 1986; Warden & Lin, 2000) show that contextual differences related to academic requirements, travel, friendship, and knowledge affect motivation. Dörnyei (1994) agrees with Crookes and Schmidt's (1991) argument that the product-oriented approach of the model does not explain student motivation in a language class. Most importantly, Ushioda (2013) argues that the concept of integrating with a target language community loses reverence in a

globalised world where English is becoming a required skill, where the target language community is not clearly known, and where internet technology dissolves boundaries between different communities. The arguments highlight the difference between individuals learning a foreign language in a foreign environment to those learning in an environment that the target language is not spoken. Gardner (2010) continues to respond to criticism of the socio-educational model; he agrees that the model has limitations but believes the socio-education model to be better than any other model.

The socio-educational model was widely accepted until the 1990s. However, since then, researchers (e.g., Brown, 1990; Julkunen, 1989; Crookes & Schmidt, 1991; Oxford & Shearin, 1994; Skehan, 1989) called for new paradigms, based on research findings from the education and psychology fields, to explore language learning motivation in relation to factors such as cognitive processing, classroom motivational techniques and activities, syllabus design, and out-of-class considerations. Rather than looking at an external reference group, researchers (e.g., Dörnyei & Csizér, 2002; Ushioda, 2013) say one's internal representation of themselves is more important; hence, language learning motivation is closely associated with the concept of self and identity.

In between the time that the socio-educational model was most prominent, and the L2MSS framework was developed, the cognitive-situated period of L2 research flourished. The cognitive-situated period viewed motivation as more complex, dynamic, and situated compared to Gardner's (1985) socio-education model (Dornyei & Ushioda, 2011). Dörnyei (2005) explains that the cognitive-situated period of L2 motivation aimed to align L2 motivation theory with mainstream psychological theory and shift the focus of research from ethnolinguistic communities of language learners to more situated and specific learning contexts such as language classrooms, building on what was learned during the social-psychological period, and searching for ways to apply theory to the classroom environment.

Cognitive theories of motivation are concerned with the "role of mental structures, beliefs and information-processing mechanisms in shaping individual behaviour and

action. Motivation is viewed as located within the individual, though naturally the individual's cognitions and perceptions may be influenced by various social and environmental factors" (Dörnyei & Ushioda, 2011, pp. 12-13). Crooke and Schmidt (1991) set the agenda for the cognitive-situated period by introducing various mainstream psychological theories to the L2 research community and suggesting ways for researchers to proceed. Oxford and Shearin (1994, 1996) also examined mainstream psychological theory and determined that needs theories, expectancy-value theories, equity theories, reinforcement theories, social cognition theories, achievement goal theory, Piaget's cognitive developmental theory, and Vygotsky's sociocultural theory can be applied to L2 motivation theory generation. The most famous and influential theory borrowed from mainstream psychology and applied to L2 motivation research was Deci and Ryan's (1985) SDT.

2.2.1 Self-Determination Theory

Self-determination theory is a comprehensive meta-theory of human motivation that has been widely influential in various academic fields such as psychology, business, health, and education as it provides a framework to understand what motivates people, how motivation can be changed, and the ramifications of motivation change (Landers et al., 2015). Dörnyei and Ushioda (2011) explain that SDT grew from studies in the 1970s comparing intrinsic and extrinsic motivation that failed to explain the complexity of human motivation through only the two-way intrinsic-extrinsic dichotomy. SDT continues to evolve to address changes in society. The fact that SDT is still being used as a theoretical framework to analyse current trends in psychology and education provides testament to its reliability and validity for understanding the human psyche (e.g., Sergis, Sampson, & Pelliccione, 2018; Vansteenkistem 2018).

Self-determination theory asserts that the more a person has control over the things they do, the more they will feel fulfilled, motivated and engaged. SDT is concerned with the type of motivation, not the amount of motivation. The two types of motivation SDT focuses on are controlled and autonomous motivation; synonymously known as extrinsic and intrinsic motivation. SDT is primarily based

on two academic theories: cognitive evaluation theory (CET) and organismic integration theory (OIT) (Ackerman, 2019).

The main premise of CET is that, if a human's psychological needs of autonomy, competence, and relatedness are fostered within a domain, they will be personally fulfilled, and intrinsically motivated to perform well and engage in that domain (Ryan & Deci, 2009). Ryan and Deci explain that autonomy refers to a sense of free will and being able to make meaningful decisions, competence refers to a feeling of being effective or a sense of accomplishment, and relatedness refers to a meaningful connection with other people. Activities that support autonomy, competence, and relatedness not only foster intrinsic motivation towards the domain, but also lead to various other higher quality learning outcomes (Carreira, 2012; Hiromori, 2003). In contrast, Ryan and Deci (2009) warn that activities or environments in which the tenets of SDT are unsupported or not supported correctly could negatively impact intrinsic motivation and lead to lower quality learning outcomes.

Organismic integration theory presents three categories of motivation individuals can have towards a domain: amotivation, extrinsic motivation, and intrinsic motivation (Ryan & Deci, 2000). The OIT divides extrinsic motivation into four different types based on associated regulatory style, perceived locus of causality, and relevant regulatory processes. Listed from most external to most internal, the four types of extrinsic motivation are: external regulation, introjected regulation, identified regulation, integrated regulation. The different types of motivation are further explained in Section 2.2.2. A central belief of the OIT continuum is that externally regulated behaviour can become internally integrated behaviour if the environment supports the psychological needs of autonomy, competence, and relatedness (Dörnyei & Ushioda, 2011).

The application of SDT to the L2/FL context developed from knowledge gained during the social-psychological period of L2 motivation research. Even though the concept of integrativeness outlined in the Gardner's (2010) socio-education model is not represented as a motivation orientation on the OIT continuum, it has been shown to overlap with the more self-determined orientations (e.g., Noels, 2001a, 2001b, 2005; Noels, Clement, & Pelletier, 2001). McEown, Noels, and Saumure (2014)

explained that “SDT orientations tended to better predict aspects of motivation related to learning engagement, whereas the integrative orientation tended to better predict motivational aspects related to cultural and community engagement” (p. 229). Noels et al. (2001) showed a strong correlation between instrumental orientation and external regulation.

Investigations into the application of SDT to the EFL context (e.g., McEown et al., 2014; Noels, 2013) show a range of positive learning outcomes associated with students who report high levels of self-determined motivation; for example, higher levels of engagement with learning activities (Ma, 2009; Spratt, Humphreys, & Chan, 2002; Tanaka, 2009) and the target language community (Comanaru & Noels, 2009); more willingness to communicate (Nishida, 2012); more frequent use of, and better L2 performance (Ehrman, 1996; Goldberg & Noels, 2006; Noels et al., 1999, 2001; Vandergrift, 2005), stronger determination to persist and higher L2 motivational intensity (Noels, 2001a; Noels et al., 1999; Ramage, 1990); better metacognitive awareness (Spratt et al., 2002; Vandergrift, 2005); less anxiety, higher positive attitudes towards language learning and elevated feelings of self-efficacy (Ehrman, 1996; Noels, 2001b; Schmidt, Boraie, & Kassabgy, 1996). The positive outcomes suggest that language learning activities should be designed to support and foster self-determined intrinsic motivation.

The three tenets of SDT provide a simple framework to design L2/FL activities that foster intrinsic motivation. Competence can be fostered through suitably challenging activities, which have clear goals, and provide constructive and informative feedback (Reeve, 2002; Ryan & Deci, 2000; Vansteenkiste, Sierens, Soenens, Lucykx, & Lens, 2009). Many studies (e.g., Deci, Vallerand, Pelletier, & Ryan, 1991; Hiromori, 2003; Otoshi & Heffernan, 2011; Vallerand, 1983) show that fostering competence increases intrinsic motivation, especially when students have a sense of autonomy. Autonomy can be fostered by allowing students the freedom to make personally meaningful choices about what learning content they interact with, and how they interact with it (McEown et al., 2014). L2/FL environments that support autonomy lead to an increase in students’ self-determined motivation (Noels, 2005; Noels et al., 1999, 2000, 2001; Vallerand, 1983), deeper engagement and better performance with the learning content (Assor, Kaplan, & Roth, 2002; DeCharms, 1984; Flink,

Boggiano, & Barrett, 1990), and increase perceived competence and self-worth (Noels et al., 2001; Reeve, 2002; Zhou, Ma, & Deci, 2009). Relatedness can be fostered through activities that are personally relevant to the students' lives, which encourage meaningful interactions with other people (McEown et al., 2014). Relatedness can increase self-determined motivation, and facilitate deeper engagement, higher achievement, and improved personal well-being (Furrer & Skinner, 2003; Hiromori, 2003; Kochanska, 2002).

2.2.2 The Language Learning Orientation Scale

The Language Learning Orientations Scale (LLOS) is a questionnaire that was developed to measure students' L2 motivational orientation from an SDT perspective (Noels et al., 2000). The LLOS is comprised of seven subscales that measure amotivation, three types of extrinsic motivation, and three types of intrinsic motivation. Each type of motivation is categorised based on its regulatory style, perceived locus of causality, and relevant regulatory processes. The LLOS subscales are similar to the OIT subscales; however, the LLOS does not measure integrated regulation, and identifies three types of intrinsic motivation rather than one.

The three types of extrinsic motivation are presented on a continuum from left-to-right depending on their perceived locus of causality; the left represents external control and the right represents internalised autonomy. On the far left of the continuum is the least autonomous, least self-determined type of extrinsic motivation, external regulation. External regulation derives from an external source such as a boss, parent, or teacher telling someone to do something; rewards or punishment are used by the external source to regulate someone's behaviour. Once the rewards and punishments are removed, it is expected that the person will cease the activity as they have no incentive. To the right of external regulation on the continuum is introjected regulation. Introjected regulation is still externally derived; however, someone has some internal motivation to complete an activity to avoid guilt or maintain or enhance their self-esteem; for example, a student studying for the TOEIC test because they would feel embarrassed if they could not get a good score.

The next type of extrinsic motivation that is more autonomous and more self-determined is identified regulation. Identified regulation occurs when a person considers the external activity to be personally important. For example, if a student feels that speaking English fluently is important, they will endure monotonous speaking drills to become fluent. A central belief of the continuum is that externally regulated behaviour can become internally integrated behaviour if the environment supports the psychological needs of autonomy, competence, and relatedness.

Unlike the three extrinsic motivation subscales, the three intrinsic motivation subscales do not differ in their levels of self-determination. The accomplishment subscale represents intrinsic motivation that fosters through the feeling of satisfaction one gets when achieving a goal such as understanding a newspaper written in the L2. The knowledge subscale represents intrinsic motivation that fosters when engaging in an activity for the pleasure and satisfaction associated with learning something new such as learning something interesting about the L2 culture. The stimulation subscale represents intrinsic motivation that fosters when someone engages in an activity that provides positive feelings such as the pleasure one might feel when listening to music in the L2 (Vallerand, 1997; Vallerand & Ratelle, 2002).

The LLOS has proven to be an effective tool for understanding the ramifications of L2 motivational orientation on L2 performance (Comanaru & Noels, 2009; Pae, 2008; Rubenfeld, Sinclair, & Clément, 2007; Vandergrift, 2005; Wu, 2003). McEown et al. (2014) make two salient points about classifying students based on their regulation: (1) learners may clearly align with one type of regulation; however, it is more likely they will possess varying degrees of each type of regulation; (2) just because one orientation is lower than another orientation, it does not mean that the lower orientation is not important; Green (1999) and McEown et al. (2014) exemplify the two points. Green (1999) found that students learning English in Hong Kong had high levels of identified regulation and external regulation. McEown et al. (2014) showed that Canadian students learning Japanese had high levels of identified regulation, intrinsic motivation, and external regulation in that order from high to low; this basically meant that learning Japanese was valuable for their personal goals, they enjoyed learning Japanese, and it was necessary to study to address some type of external pressure such as passing a course.

2.2.3 Foreign Language Motivation in the Japanese Context

The Japanese people's struggle with the English language is a well-known issue. In 2017, Japanese students who took the Test of English as a Foreign Language (TOEFL) internet-based examination ranked 27 out of the 29 Asian countries (ETS, 2017). In a worldwide comparison of non-English speaking countries, EF English Proficiency Index (EF EPI, 2018) assigned Japan a *low* ranking for English proficiency; low is the fourth lowest ranking out of the five possible rankings. Ryan (2009) says there is "a permanent sense of crisis" (p. 407) surrounding English education in Japan. In contrast to Japan's poor English performance, the Japanese economy is the third largest in the world (Kennedy, 2018). This disparity between Japan's economic position and their English language ability is of concern to the Japanese government in an ever-increasing competitive globalised world in which the English language is the lingua franca (Kubota, 1998). This led researchers to explore the problematic Japanese context in which they determined that the poor performance can be attributed to low intrinsic FL motivation to learn EFL (e.g., McVeigh, 2004; Nakata, 2006; Ushioda, 2013). This section uses SDT in conjunction with other relevant theories to explain the three main factors affecting FL motivation in Japan FL, and to provide a rationale for gamification to be explored as a pedagogical tool that supports FL motivation and FL acquisition.

The first main factor affecting intrinsic FL motivation in the Japanese context is related to Japan's geographical location. The literature (e.g., LoCastro, 2001; Ushioda, 2013) explains that because English language ability is not required in Japan, and because Japanese people are rarely exposed to foreigners, Japanese people have no pressing need to become proficient EFL communicators. From an SDT perspective (e.g., Deci & Ryan, 1985; Noels et al., 2000), Japan's geographical and cultural isolation could have resulted in a large portion of the Japanese population feeling amotivation towards learning EFL. To counter the low L2/FL intrinsic motivation derived from geographical isolation, Yashima (2002, 2009) suggests that language learning activities should strive to get students interested in the English-speaking world outside of Japan; this construct is referred to as international posture. The findings from Aubrey and Nowlan (2013) support the fostering of international posture as an approach to increase L2/FL intrinsic motivation.

The second main factor affecting intrinsic FL motivation in the Japanese context is the Japanese education system, specifically, the exam-based culture and the compulsory nature of English classes (Nagano, 2009; Tsuchiya, 2006). Japan is a society that places a high emphasis on high school and university entrance exams; future employment is strongly affected by what learning institutions someone attends (Nagano, 2009). The entrance exams are described as a "national obsession" (Rohlen, 1983, p. 77) that foster fierce competition (Amano, 1996) which result in "exam hell" (Ushioda, 2013, p. 5) for partaking students. Testing students' English ability is an important component of the entrance exams; however, the English component has been criticised for poor design that tests discrete items of grammatical competence rather than communicative ability (Brown, 2002; Brown & Yamashita, 1995; McVeigh, 2001). This results in teachers using antiquated pedagogical approaches, such as the grammar-translation method, that inadvertently over-emphasises the importance of test-performance over communicative ability (Gorsuch, 1998; Taguchi; 2005; Ryan, 2009). This leads students to believe that the most important thing about learning English is being able to remember the answers to test questions.

The university entrance exam is very important, and it is considered the final important exam students need to take; once students complete it, motivation to continue studying declines (Ushioda, 2013). In contrast to *exam hell*, the university experience is described as a "leisure-land" (Clark, 2010), and *4-year moratorium* (Sugimoto, 2010), a place where student motivation to study declines due to a perception that the hard-work was complete when university acceptance was confirmed. Berwick and Ross (1989) describe English education in Japanese universities as a *motivational wasteland* and they believe that students' motivational intensity to study English peaks in the final year of high school. However, as English is a mandatory subject for most university courses at Japanese universities, teachers are faced with a situation in which they might be teaching students who lack L2 intrinsic motivation.

Self-determination theory related studies (e.g., Hamada & Kito, 2007; Ockert, 2011; Tachibana, Matsukawa, & Zhong, 1996) show that Japanese students generally

possess stronger extrinsic orientations of motivation than intrinsic orientations. SDT suggests that, because parents and teachers demand compliance through rewards and punishments, this decreases students' autonomy and increases their external regulation towards learning EFL. Once students complete their entrance exams, the external regulation from parents and teachers declines, and so does their motivation to study English; however, this is not true for all students.

Kimura, Nakata, and Okumura (2001) found that L2/FL motivation orientation depends on academic level and major. Yashima et al. (2009) found that Japanese learners of English can have strong identified regulation, and students who are comfortable interacting with native speakers of English are more likely to be intrinsically motivated to learn English. The literature (e.g. Abe, Shimizu, Okuda, Ishizuka, & Ueda, 2010; Berwick & Ross, 1989; O'Donnell, 2003) also shows that many Japanese students do not have clear language goals; goal-setting is important for L2/FL motivation (Dörnyei & Ushioda, 2011). Lack of clear goals leads to students who are extrinsically or instrumentally oriented, rather than intrinsically motivated (Ockert, 2011).

The third main factor affecting intrinsic FL motivation in the Japanese context is students' high anxiety and low self-efficacy (Andrade & Williams, 2009; Ebata, 2008). Krashen's (1982) affective filter hypothesis postulates that language learners with high motivation, high self-confidence, and low anxiety are ideally equipped to acquire the L2; students who do not possess those characteristics are less likely to be successful. Studies (e.g., Carreira, 2006; Matsuda & Gobel, 2004; Yamashiro & McLaughlin, 2001; Yashima et al., 2009) show that too much anxiety affects intrinsic motivation in negative ways. Andrade and Williams (2009) found that 75% of Japanese students studying EFL at university are affected by anxiety; this stems from and leads to a reduction in self-confidence to learn and use EFL (Hashimoto, 2002).

The literature (e.g., Burden, 2002; Ebata, 2008; Kobayashi & Brown, 2003; O'Donnell, 2003; Woodrow, 2006) shows that Japanese foreign language learners lack self-confidence in their ability to use foreign languages. The literature (e.g., Benson, 1991; Burden, 2002; Falout, 2004; Kikuchi & Sakai, 2009; Matsuda &

Gobel, 2004; Tsuchiya, 2006) also attributes the lack of communicative confidence to Japan being a country that lacks opportunities to use English in meaningful ways, and to the education system that uses outdated pedagogical approaches that emphasise test performance over communicative ability. Learning environments which foster cooperativeness can reduce communication-related anxiety, which leads to increased self-confidence and motivation, especially for students with low motivation and high anxiety (Koga, 2010).

From around the year 2000, the cognitive-situated period of L2 research merged into the process-oriented period of research (Dörnyei, 2005). The process-oriented period of L2 research explores how L2 motivation changes over time "at either the micro level (e.g. task motivation) or the more macro level (e.g. during a course of study, over a person's learning history or across the lifespan)" (Dörnyei & Ushioda, 2011, p. 60). As studies (e.g., Guilloteaux & Dörnyei, 2008; Sugita & Takeuchi, 2010, Sugita McEown, & Takeuchi, 2014) continue to show a positive correlation between teachers' using motivational strategies and students' motivation to learn, researchers (e.g., McEown et al., 2014) continue to look for practical ways that motivation theory research can inform pedagogy. Pedagogical approaches have evolved from the grammar-translation method to more effective approaches such as autonomous learning (e.g., Allwright, 1990; Dam, 1995; Gremmo & Riley, 1995; Little, 1991) and task-based language teaching (TBLT) (e.g., Willis & Willis, 2007) that are now being further leveraged through various computer-assisted language learning (CALL) applications (e.g., Warschauer, 1996).

Gamification is a new tool to support EFL pedagogy that combines aspects of various pedagogical approaches with gamification components with the intention to increase student engagement and L2/FL motivation. Simple gamification implementations rely on extrinsic rewards to influence a target behaviour; more sophisticated gamification implementations can potentially leverage motivation theory such as SDT to increase students' L2/FL intrinsic motivation. Gamification has been successfully used in the Japanese business sector to motivate consumer and employee behaviour (e.g., Hooghiemstra, 2017; Werbach & Hunter, 2012); however, it remains largely unexplored in a Japanese EFL context (explained in Baber, 2015). This thesis examines the application of gamification to the structure of an EFL

course at a Japanese university to analyse how it affects student performance and FL motivation.

2.3 Gamification

Gamification is the concept of applying game elements and game design techniques to non-game contexts to motivate behaviour (Deterding et al., 2011; Werbach & Hunter, 2012). Motivating behaviour using techniques commonly found in games has a long history: from the Boy Scouts of America awarding merit badges to encourage members to learn new skills in 1911 (Growth Engineering, 2016), to Cracker Jack putting surprise toys in every box of popcorn in 1912 (Werbach & Hunter, 2012), to American Airlines launching the first frequent flyer program in 1981. Companies and educational institutions have been looking to games (e.g., Coonradt, 2012; Malone, 1980) for innovative ways to motivate their customers, employees, and students for a long time. After the term gamification was adopted in 2010 (Fuchs, Fizek, Ruffino, & Schrape, 2014; Hamari et al., 2014) to describe the application of game techniques to non-game contexts, interest in gamification rose due to five catalysts: Speed Camera Lottery, Nike+ (explained in Section 1.1), Farmville, Jane McGonigal's (2011) book *Reality is broken*, and Lee Sheldon's (2011) book *The Multiplayer Classroom*.

Speed Camera Lottery was a system designed to encourage drivers to drive within the speed limit. It was tested in Stockholm, Sweden in 2010. A speed camera was placed on the side of a road to measure the speed of passing drivers. Each driver's speed was displayed on a digital display board that was located just ahead of the camera. If a driver were caught speeding, they would have to pay a speeding fine as usual. However, drivers who were not speeding were automatically entered into a Lotto which gave them a chance to win some of the money collected from the speeding fines. During the three-day trial period, the average speed dropped by 20% (Chou, 2015). This example of gamification used game mechanics such as feedback and rewards to encourage behaviour change. Speed Camera Lottery can be classified as a type of social behavioural change gamification that aims to bridge the gap between current behaviour and desired behaviour.

In 2010, *Farmville*, a social farm-management game, had 83 million active Facebook players, according to Thulin (2018). Pioneering gamification proponents were intrigued by Farmville's ability to engage and motivate players in the seemingly monotonous activity of virtual farming. McGonigal (2011) asserts that the reason for this game's popularity was due to the combination of gamification techniques such as points, badges, and leaderboards that allowed players to enjoy their productivity in a socially connected environment. Gamification proponents (e.g., Werbach & Hunter, 2012) hypothesised that the same techniques used in Farmville could be used to motivate people in various other situations. However, there was one caveat: gamification implementations need to be more than just adding points, badges, and leaderboards to an activity if long-term sustained engagement is the goal.

McGonigal (2011), in her book *Reality is broken: Why games make us better and how they can change the world* asserts that peoples' lives can be improved, and the world can become a better place if we take what we know about games and apply that knowledge to real-life situations. McGonigal wrote about serious games that raise awareness and attempt to solve complex world problems. For example, *Darfur is Dying* was a game released in 2006 to raise awareness about the humanitarian crisis in Darfur, Sudan. She also wrote about alternate reality games that aim to motivate behaviour by placing real-world goals within a fictional setting; for example, *Chore Wars* was developed to motivate people who lived in a shared environment to do more chores. Davis (2007) explains that the system tracks the number of chores someone does and rewards them experience points (XP) and virtual gold when a chore has been completed. Players decide how the virtual gold can be redeemed in real life.

The final catalyst, specific to the context of this thesis, was Lee Sheldon's (2011) seminal book *The Multiplayer Classroom* in which he introduced the concept of gamified instructional design. Gamified instructional design is a type of behavioural change gamification that aims to engage students in an educational context through the application of game-design techniques. Gamified instructional design can occur on a macro level in which a whole class is gamified (e.g., Sheldon, 2011), or on a micro level in which specific learning activities are gamified (e.g., duolingo.com).

When a whole class is gamified, the teacher redesigns various elements of the class to be more game-like. For example, instead of students completing homework assignments, they complete quests. Instead of the teacher giving students a grade for completing a quest, the teacher awards (XP) which gets displayed on a class leaderboard. As students accumulate XP they level-up; at the end of the course, the level a student attains could be converted to a class grade. Students could receive XP for a wide range of things such as completing homework, attending class, and any other type of activity the teacher would like to reward. This thesis focuses on gamified instructional design.

2.3.1 Gamification Design Framework

Sheldon's (2011) work laid the foundation for gamified instructional design, but since then, other academics have provided more explicit frameworks to guide gamification design; for example, the MDA (Hunicke, Leblanc & Zubek, 2004) and Octalysis (Chou, 2015) frameworks. Borrowed from traditional game design, there are over 100 different game mechanics or game elements that could be used for gamification implementations (Dubravac, 2012). A framework allows gamification designers to make informed decisions about which components to include in a gamification implementation. The frameworks vary in their complexity and appropriateness for a situation.

This study uses Werbach and Hunter's (2012) The Pyramid of Elements framework to guide the development of meaningful gamified instructional design for an EFL course. The framework was utilised as it focuses on the cause and effect of different gamification elements. Apart from the aesthetics, Werbach and Hunter (2012) explain that there are three major design elements to consider when gamifying an activity: dynamics, mechanics, and components. Werbach and Hunter describe dynamics, mechanics, and components as elements which should be conceptualised as a pyramid structure. Dynamics are at the highest level of the pyramid, mechanics are in the middle, and the components are the base of the pyramid. Components are the 'things' (e.g., points, levels, achievements, content unlocking, leaderboards, quests, social media); mechanics are the specified interactions between the 'things',

the processes that drive the action (e.g., competition, cooperation, feedback, rewards, challenge, win-state); and dynamics are the higher-level emergent interactions which are a result of the utilised components and mechanics (e.g., personal progression, relationship building, emotional reaction).

A dynamic cannot be guaranteed but, through the correct component and mechanic implementation, the likelihood of success increases. The components, mechanics, and dynamics can overlap and have multiple connections to themselves and the other levels. For example, two students (component-teams) working together (mechanic-cooperation) on a learning task (component-quest) perform better than the other teams (mechanic-competition). They are rewarded (mechanic-reward) with a badge (component-badge) which makes them feel proud (dynamic-emotion), builds their relationship (dynamic-relationship), and motivates them to continue studying (dynamic-progression).

Werbach and Hunter (2012) suggest looking to psychological theories to support the design of the gamification framework. Cognitive theories such as social comparison theory (Buunk & Gibbons, 2006) and goal-setting theory (Locke & Latham, 2006) can provide useful insight; however, it is SDT (Deci & Ryan, 1985) that is most commonly associated with gamification (Chou, 2015; Werbach & Hunter, 2012). A possible reason that gamification is commonly associated with SDT could be due to the wide range of extrinsic and intrinsic motivators that are incorporated into gamification implementations.

Self-determination theory can guide the design to either reward-based gamification or meaningful gamification (Tan & Hew, 2016). Extrinsic reward-based gamification focuses on the application of extrinsic rewards to motivate behaviour. Meaningful gamification focuses on the application of psychological theory to intrinsically motivate behaviour while using extrinsic rewards to direct behaviour. The approach taken depends on the goals of the implementation. If the goal is intense short-term engagement, extrinsic reward-based gamification may be suitable; if the goal is long-term engagement, meaningful gamification is more appropriate. This study is a meaningful gamification implementation as extrinsic rewards will be used to

motivate short-term behaviour, but the more important aim is to foster long-term intrinsic FL motivation.

2.3.2 Gamification and CALL

The recent boom of gamification has come from improvements in digital technology being able to efficiently present, track, share, and provide novel feedback about data. Even though digital technology is not required for gamified instructional design, it allows for various gamification processes to happen efficiently and semi-automatically (Garland, 2015). The application of gamification to an EFL course can look to the field of computer-assisted language learning (CALL) for guidance. CALL is "an overarching term encompassing an ever-growing range of applications" (Son & Windeatt, 2017, p. 3) that support the language learning process. CALL technology can offer a range of advantages over traditional approaches; for example, efficient and accurate content delivery, expedited feedback on completed work, and the ability to easily track, access, and present data about students' performance (Devers & Gurung, 2015; Dina & Ciornei, 2013). The similarities and successes of the recent gamification boom and CALL suggest that they are uniquely positioned to work together to create an effective gamified EFL course. The application of CALL to a gamified EFL course can be viewed from two perspectives: learning outcomes and class management (Barata, 2017; Hanus & Fox, 2015).

The learning outcomes perspective of a gamified CALL environment is concerned with how digital technology can support FL acquisition and FL motivation. Two ways in which CALL can support FL acquisition and FL motivation are through automated and accurate content delivery, and content choice. Krashen's (2009) Input Hypothesis posits that L2 acquisition occurs when learners understand the language content being delivered to them with the content being slightly above their current ability (i+1). Therefore, based on the assumption that the Input Hypothesis is a valid construct, if the gamified system supports the delivery of learning content that is i+1, or if the system allows students to choose content that is i+1, the system can theoretically support FL acquisition.

The autonomous learning literature (e.g., Benson, 2007; Pemberton, Toogood, & Barfield, 2009) explains that allowing and encouraging students to make choices about the learning content used as part of the course fosters intrinsic motivation. SDT (Ryan & Deci, 2009) also suggests that intrinsic motivation fosters when students make meaningful choices about their learning content, and when students complete learning activities that are suitably challenging. Therefore, if the gamified system allows students to choose learning content that is suitably challenging and personally engaging, it can potentially support FL motivation. If students have low motivation to do learning activities that support FL acquisition and FL motivation growth, gamification components, such as progress bars or leaderboards, might be used to encourage a performance increase. It should be noted that some studies (e.g., Devers & Gurung, 2015; Fried, 2008; Goodwin, 2011; Higgins, Beauchamp, & Miller, 2007; Marklein, 2010) warn that if CALL activities are poorly designed or lead to unwanted distractions, they may not have a positive effect on learning.

The Clark-Kozma debate is a long-standing debate about the application of technology to learning. In 1983, Clark (1983) argued that "media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition" (p. 445). Clark (1994) asserted that not only does technology "not influence learning, but it will never influence learning, and that media is neither sufficient for nor necessary to learning" (p. 23). After 30 years, Clark (2012) still maintains his position by saying that "there is strong evidence that many very different media attributes accomplish the same learning goal" (p. 175). Kozma (1994, 2000), on the other hand, says that media will have an effect on learning and that, if more research is done, we will be able to prove this. Ultimately, what can be learned from this debate is the importance of context. CALL activities are not necessarily better than non-CALL activities; their efficacy depends on how they are utilised, and their suitability for the learning context. Teachers need to decide if a CALL activity is a suitable delivery truck.

The class management perspective of a gamified CALL environment is concerned with how digital technology can efficiently handle various class management aspects of a gamified course. For example, digital technology can be used to create the

ecosystem for the gamified course to exist. Various websites and digital tools can aggregate data from multiple sources and then create gamification components such as leaderboards and progress bars. The digital ecosystem can automate and centralise a range of other class management processes such as content storage and delivery, class communications, feedback, and assessment.

Some websites have been developed for the purpose of gamifying academic courses; for example, Rezzly (https://www.rezzly.com) and Classcraft (https://www.classcraft.com) are gamified learning management systems that provide an online platform for teachers to build, deliver, and track learning content while being supported by gamification components such as narratives, avatars, leaderboards, points, and levels. Rezzly and Classcraft may be suitable platforms for teachers to use to achieve their gamified goals; however, if they are not suitable, teachers can combine other software and websites to develop a more suitable platform. Section 3.4 describes how various software solutions were combined to create a suitable gamification platform for this study.

2.3.3 Gamified Instructional Design Research

Studies examining the application of gamification to foreign language courses are limited (Reinhardt, 2019). However, gamification research conducted in other disciplines can guide the design of future gamification implementations and identify salient gaps in the literature that need to be addressed. Gamified instructional design literature begins with Sheldon's (2011) seminal book which includes eight case studies that describe and discuss how gamification was implemented in classes from primary school to university in various courses such as biology, history, mathematics, technology, and general education. Even though the research methodologies in Sheldon's case studies sometimes lack academic rigour, the teachers' observations and opinions provide exploratory insight and guidance which other researchers have followed on from. More recent studies have been conducted which use more reliable research methodologies (e.g., Hanus & Fox, 2015), with larger cohorts of participants (e.g., Poondej & Lerdpornkulrat, 2016), over longer periods of time (e.g., Barata et al., 2017), and based on more informed gamification

design (e.g., Nicholson, 2013). This section provides a general overview of the gamified instructional design research which has prompted the context of this thesis.

An underlining theme running through the literature is the importance of context (Buckley et al., 2017; Garland, 2015). There are many contextual specificities which can affect the efficacy of a gamified course. Different students respond differently to the same gamification components (Barata et al., 2017). For example, the study by Koivisto and Hamari (2014) determined that, even though men find traditional games more playful than women, females in their study found the gamified experience more playful than men. Goehle (2013) found that students who play video games were more motivated by gamification components than students who did not play video games. Poondej and Lerdpornkulrat (2016) examined individual differences and found that students with higher computer literacy skills were more engaged by a gamified course than students who were not as comfortable with computers. Aldemir et al. (2018) showed that some students like the competitive environment of a gamified class, whereas other students prefer a collaborative environment.

One of the initial goals of this thesis was to explore the use of gamification to increase the short-term performance of EFL students. Multiple studies (e.g., Charles, Charles, McNeill, Bustard, & Black, 2011; Li et al., 2013; Mekler et al., 2013) show that gamification has been able to increase students' performance, engagement, participation, and attendance. However, other studies (e.g., Goehle, 2013) did not find evidence that gamification affected performance in a positive or negative way. When performance has increased, there is concern about the quality of the perceived performance increase. The students may just be performing in order to achieve the extrinsic goals of the gamified course, rather than performing to learn. Some studies (e.g., Barata et al., 2013) show that gamification has led to increased class grades; however, other studies (e.g., de-Marcos et al., 2014; Hanus & Fox, 2015) found that students in a gamified course received lower exam scores at the end of the course. Due to concerns about short-term performance gains coming at the cost of long-term engagement towards the learning domain (Garland, 2015; Hamari et al., 2014), researchers have examined the effect of gamification on motivation.

Many studies (e.g., Abramovich et al., 2013; Barata et al., 2013; De-Marcos et al.,

2014; Dominguez et al., 2013; Goehle, 2013; Li et al., 2013; McDaniel, Lindgren, & Friskics, 2012; O'Donovan, Gain & Marais, 2013) have reported that gamification positively affects student motivation. Charles et al. (2011) and Goehle (2013) attributed the rise in motivation to the gamification fostering feelings of acknowledgement and accomplishment. Some studies (e.g., De Schutter & Abeele, 2014) did not conclude that gamification led to higher levels of intrinsic motivation or engagement compared to traditional teaching techniques, and other studies (e.g., Berkling & Thomas, 2013; Dominguez et al., 2013; Haaranen, Ihantola, Hakulinen, and Korhonen, 2014; Hanus & Fox, 2015; Meyer, 2008) reported a mixed, neutral, or negative impact on motivation. Hanus and Fox (2015) determined that students' motivation inclination affects how they respond to gamification. De Schutter and Abeele (2014) found that the role of the teacher and the quality of the teaching materials were more important for facilitating intrinsic motivation than the gamification itself.

The literature presented in this section was generally positive but sometimes mixed towards gamified instructional design; however, research deficiencies associated with small sample sizes (e.g., Abramovich et al, 2013; Meyer, 2008; O'Donovan et al., 2013), lack of control groups (e.g., Sheldon, 2011), short duration times of data collection (e.g., Lister, 2015), contextual differences in different gamification implementations (e.g., De Schutter & Abeele, 2014), and the unfocused approach of analysing gamification as a whole entity rather than focusing on specific gamification components (e.g., Landers et al., 2015, Lister, 2015; Sailer, Hense, Mayr, & Mandl, 2017) makes it difficult to make definitive claims about the effectiveness of specific gamification components.

Even though Çakıroglu et al. (2017) suggested that a combination of gamification components and mechanics may lead to the greatest performance and motivation outcomes, research is first required that isolates specific components. Therefore, this thesis focuses on two gamification components, leaderboards and quests. Leaderboards are examined due to their potential to increase student performance, and due to concerns that the performance gains will come at the cost of intrinsic FL motivation (Bielik, 2012). Quests are examined due to their potential to positively impact intrinsic FL motivation (Sheldon, 2012). As explained in Section 2.2.3,

motivation is an important aspect of EFL education in Japan (Nakata, 2006; Ushioda, 2013), it is important to explore the possible positive and negative ways gamification can affect it.

2.4 Leaderboards in Gamified Instructional Design

Leaderboards are a gamification component that have the potential to bridge the gap between current student behaviour and desired behaviour (Werbach & Hunter, 2012). Leaderboards are intertwined with many gamification components (e.g., points, levels), mechanics (e.g., competition, feedback), and dynamics (e.g., performance, motivation, emotions). Leaderboards can have a strong effect on the dynamics of the gamification implementation (Deterding, 2013; Sheldon, 2011). Leaderboards are one of the most common methods to gamify a course (Dicheva, Dichev, Agre, & Angelova, 2015). Their usefulness and relative ease of setting up could explain why they are so commonly used in gamification implementations (Cheong, Cheong, & Filippou, 2013; Domínguez et al., 2013; Hamari et al., 2014).

Leaderboards provide socially-comparative feedback to the users of a gamified system about their performance (Codish & Ravid, 2014); the feedback is given through points and rank. Leaderboard rank is “based on a set of criteria that is influenced by the users' behaviors towards the desired actions” (Chou, 2015, p. 121) of the system. Leaderboards aim to encourage ideal behaviour through a positive representation of a user’s points and leaderboard rank and discourage non-compliant behaviour through a negative representation of a user’s points and leaderboard rank (Kapp, 2012; Malone, 1980).

Literature that examines leaderboards in gamified EFL courses is lacking; however, leaderboard research that has occurred in other disciplines provides guidance. Investigations into the use of leaderboards has occurred in a variety of courses such as medical (e.g., Halan, Rossen, Cendan, & Lok, 2010), research methods (e.g.,Tan & Hew, 2016), information technology (e.g., Cheong et al., 2013; Domínguez et at., 2013), software development (e.g., Borys & Laskowsky, 2013; Codish & Ravid, 2014), a MOOC (e.g., Morales, Amado-Salvatierra, Hernández, Pirker, & Gütl,

2016), and EFL (e.g., Philpott, 2015a). This section examines leaderboard-related literature to provide insight into how leaderboards affect performance and motivation.

2.4.1 The Effect of Leaderboards on Performance

Studies that examine leaderboards in gamified courses show that leaderboards can positively impact learner behaviour related to performance (e.g., Aldemir et al., 2018), engagement (e.g., Barata, 2017; Poondej & Lerdpornkulrat, 2016), the amount of work students complete (e.g., Borys & Laskowsky, 2013; Domínguez et al., 2013; Halan, Rossen, Cendan, & Lok, 2010; Tan & Hew, 2016), time-on-task (e.g., Landers & Landers, 2014), maintaining performance (e.g., Mekler et al., 2013), and class attendance (e.g., Borys & Laskowsky, 2013; Krause, Mogalle, Pohl, & Williams, 2015; Morales et al., 2016); the positive impact can lead to increased academic performance (e.g., Charsky, 2010; Connolly, Boyle, Macarthur, Hainey, & Boyle, 2012; Iosup & Epema, 2014; Morales et al., 2016). The increases in performance have been attributed to an increased cognisance of the structure of the course, clear goals, the ability to self-assess performance, and social comparison influencing behaviour (Aldemir et al., 2018; Domínguez et al., 2013; Iosup & Epema, 2014; Tan & Hew, 2016). The positive findings come with three caveats that are of importance to gamifying an EFL course using leaderboards.

The first caveat is that the extrinsic rewards provided by a leaderboard may not be able to sustain engagement over an extended period-of-time, and they might limit behaviour. There are various psychological theories that provide warnings about leaderboards. Gamification researchers (e.g., Werbach & Hunter, 2012) often refer to behaviourism for a baseline understanding of the effects of extrinsic rewards on performance and motivation.

Behaviourism is an early theory of learning that is based on the idea that human and animal behaviour can be trained using conditioning stimuli (Chen, 2003). The most famous behavioural psychology studies were conducted by Ivan Pavlov (1902) and B.F. Skinner (1948). Pavlov found that dogs could be conditioned to salivate

(conditioned response) when they heard a bell ring (conditioned stimulus) as they associated that event with being fed (unconditioned stimulus) (Psychologist, n.d.). Based on Pavlov's observations, John Watson (1913) entitled this concept *classic conditioning* and claimed that it could explain all facets of human psychology. As Watson failed to understand the importance of the mind or consciousness for explaining behaviour, Skinner's (1948) follow-up work was based on the view that classic conditioning was a too simple explanation of human behaviour and the best way to understand human behaviour was to examine the causes of an action and its resulting consequences.

The term *operant conditioning* was coined by Skinner (1938) to describe the concept of altering behaviour through reinforcement. Skinner (1948) studied operant conditioning using his infamous Skinner Box and showed how positive and negative reinforcement can be used to control the behaviour of rats. The rats were quickly conditioned to either press a lever to receive food or press a lever to stop discomfort caused by an electrical current. These studies (i.e., Pavlov, 1902; Skinner, 1948) examined the differential effects of reward and punishment on learning, and the findings were extrapolated to human behaviour (Werbach & Hunter, 2012).

Behaviourists posited that, because animals and humans respond in predictable ways to external stimuli such as a reward or punishment, these reinforcements could be used to modify human behaviour (Werbach & Hunter, 2012). Ferster and Skinner (1957) determined that different scheduling patterns of reinforcement have different effects on the speed of learning (response rate) and the speed of someone quitting an activity (extinction rate). They described five different reinforcement schedules: continuous, fixed-ratio, fixed-interval, variable-ratio, variable-interval. Variable-ratio reinforcement resulted in the slowest extinction rate, and continuous reinforcement led to the quickest extinction rate. Variable-ratio reinforcement has been used to explain the addictiveness of slot machines (Weinschenk, 2013) with studies (e.g., Linnet, Møller, Peterson, Gjedde, & Doudet, 2011) into the neurochemical dopamine supporting the assertion. A behaviourist perspective of leaderboards views students as a type of 'black box' that can be manipulated using extrinsic rewards to positively and negatively reinforce behaviour (Antin & Churchill, 2011; Evans, Jennings, & Andreen, 2011).

Behaviourism studies (e.g., Ferster & Skinner, 1957) that examine how different reward reinforcement schedules affect sustained engagement suggest that humans eventually lose interest in extrinsic rewards. Frederick and Loewenstein (1999) explain that humans are also susceptible to hedonic adaption; devaluing extrinsic rewards when received, and then requiring bigger and better rewards to motivate their behaviour. Therefore, if the leaderboard is the main driver of behaviour, it seems like that students will eventually lose interest in it.

Self-determination theory also suggests that, if leaderboards rely on extrinsic rewards to motivate student behaviour, leaderboards will only be able to sustain performance for as long as the behaviour is being rewarded. Kapp (2012) warns that rewards can **"**create an artificial ceiling for performance at the rewards threshold. Once players have earned the reward, they are unlikely to continue on with the task that they were persuaded to do" (p. 221). This was true in one of Sheldon's (2011) case studies where students would 'take a break' because they had reached their goal. Other gamification studies (e.g., Borys & Laskowsky, 2013; Koivisto & Hamari, 2014; Krause et al., 2015; Mollick & Rothbard, 2014) support the notion that leaderboards may not able to sustain engagement and they may limit behaviour.

The second caveat that is important for the EFL context is that the performance represented on a leaderboard may not reflect learning has occurred. The literature (i.e., Domínguez et al., 2013; Halan, Rossen, Cendan, & Lok, 2010; Kapp, 2012; Philpott, 2015a; Tan & Hew, 2016; Werbach & Hunter, 2012) warns that, if students focus on achieving the extrinsic goals rewarded by a leaderboard, rather than the actual task, it will not lead to learning, could distract the student from what is actually important, and can result in lower quality of work. The literature (e.g., Aldemir et al., 2018; Hanus & Fox, 2015) suggests that leaderboards need to reward the right type of behaviour, not just behaviour. In contrast to the simple application of gamification to repetitive tasks (Lopez, 2011), gamifying an EFL course needs to account for L2/FL acquistion theories (Krashen, 2009; Swain, 1985) that require students to be mentally engaged with the learning content.

The third caveat is that not all students will respond well to the leaderboard, and one

aspect that affects students differently is the competitve social ranking. Leaderboards in a gamified course allow students to compare their performance to other students without others knowing “they are engaged in such deep social comparison” (Hanus & Fox, 2015, p. 154). Competition is one of the central mechanics fostered through leaderboards that gamification uses to drive behaviour. Competition can increase performance, motivation, and enjoyment of tasks, but can also lead to a range of negative outcomes such as disengagement or an unwanted feeling of pressure (Burguillo, 2010; Hakulinen, Auvinen, & Korhonen 2013; Orosz, Farkas, & Roland-Lévy, 2013; Lam, Yim, Law, & Cheung, 2004; Reeve & Deci, 1996).

Ranking students based on performance has a long history in education systems around the world. However, there are differences between a traditional academic ranking system and a gamified ranking system. First, the main purpose of a gamified ranking system is to motivate current student behaviour through competition (Buckley & Doyle, 2014), whereas the purpose of a traditional ranking system is to retrospectively assess student performance by assigning grades. Second, a gamified ranking system is constantly being updated and can be viewed by the students in some way (Hanus & Fox, 2015)., whereas a traditional ranking system may only be created at the end of the course and the students may never actually know their ranking in comparison to the other students.

Students generally respond positively to competitive leaderboards that show their rank. For example, Tran and Zeckhauser (2012) found that telling Vietnamese students enrolled in an ESL course their class ranking for practice tests led to better exams results; when the rankings were made public, the performance increased even more. Aldemir et al. (2018) and Çakıroglu et al. (2017) used a top 10 and a top 5 public leaderboard respectively and found that most students enjoyed the competitive leaderboard and their desire to gain reputation by being placed on the leaderboard led to a performance increase; however, not all students responded well to the social competition. The literature (e.g., Aldemir et al., 2018; Buckley et al., 2017; Çakıroglu et al., 2017; Heeter, Lee, Medler, & Magerko, 2011; Tan & Hew 2016; Tran & Zeckhauser 2012; Werbach & Hunter, 2012) suggests that the competition fostered through leaderboard rank is more likely to beneficial and appealing to the high performing students who like competition; performance could refer to the

students' leaderboard rank or their academic ability in the content domain.

Low performing students are more likely to respond negatively to a leaderboard that shows their rank (Aldemir et al., 2018; Buckley et al., 2017; Çakıroglu et al., 2017; Tan & Hew 2016; Werbach & Hunter, 2012). Social comparison theory (Buunk & Gibbons, 2006) suggests that when low performing students compare their leaderboard rank to higher ranked students, this negatively affects their self-esteem, which negatively affects their attitude towards the leaderboard, which negatively affects their performance, which results in a continued low leaderboard rank. Fotaris, Mastoras, Leinfellner and Rosunally (2016) exemplify the negative cycle by showing that student engagement declined as their leaderboard ranking declined.

Only one study, Bursztyn and Jensen (2015), showed that leaderboards led to a decrease in performance for high performing students. The study examined the use of a leaderboard that showed only the top 3 performers in computer-based high school remedial courses. The results of the study showed a 24% decline in performance overall, and a 40% decline in performance for the students most at risk of appearing on the leaderboard. The researchers attribute results with students wanting to avoid negative feelings associated with being singled out within their peer group. The competitive social ranking system of a leaderboard may negatively impact the performance of low performing students or those who have negative perceptions of the leaderboard.

2.4.2 The Effect of Leaderboards on Motivation

There is justified concern about the negative effects an extrinsically rewarding leaderboard can have on students' L2/FL intrinsic motivation (Bielik, 2012; Deci et al., 1999). Extrinsically rewarding people for previously unrewarded activity or activities they are already intrinsically motivated to do can lead to negative engagement or poorer-quality work (Hanus & Fox, 2015; McGonigal, 2011; Werbach & Hunter, 2012). The shift to extrinsic motivation undermines intrinsic motivation and can result in disinterest towards the activity once the rewards are removed if the prior intrinsic motivation does not return (Kohn, 1999; Nicholson,

2015; Tang & Hall, 1995). This concept is referred to as the overjustification effect, and it directly challenges Skinner's (1938) operant conditioning theory in relation to the importance of reinforcement. The overjustification effect has shown to be true in many settings from children drawing, blood donation, teachers' salaries, and puzzle solving (Werbach & Hunter, 2012). However, some researchers (e.g., Cameron, 2001; Cameron & David Pierce, 1994; Lepper, Greene, & Nisbett, 1973) have rejected the overjustification effect as not true, or only sometimes true in limited situations.

The cognitive evaluation theory aspect of SDT suggests that the rewards used by a leaderboard negatively impact intrinsic motivation if they are perceived as controlling. Controlling rewards lead to feelings of powerless and incompetence, resulting in a decrease of intrinsic motivation (Hanus & Fox, 2015). Leaderboards are often perceived by students as instruments of control, regardless of rank (Hanus & Fox, 2015; Philpott, 2015a). Philpott (2015a) found that when students in a gamified EFL course saw their leaderboard ranking, they felt motivated to do more of the activity that the leaderboard was rewarding; the low and middle ranked students felt more pressure to perform better, whereas students at the top of the leaderboard felt more motivated to maintain their rank. Hanus and Fox (2015), on the other hand, found that, compared to the non-gamified version of a course that did not include a leaderboard, students in the gamified course showed less intrinsic motivation, satisfaction, and empowerment as the course progressed, which correlated with lower final exam scores.

Deci et al.'s (1999) meta-analysis of 128 studies that examined the effects of extrinsic rewards on intrinsic motivation found that tangible rewards that were engagement-contingent, completion-contingent, or performance-contingent undermine intrinsic motivation. The literature (e.g., Dominguez et al., 2013; Hanus & Fox, 2015) shows that leaderboards negatively impact the intrinsic motivation of students who do not like competition. Therefore, a leaderboard in a gamified class could potentially negatively affect the intrinsic FL motivation of students who are already intrinsically motivated to do the activities the leaderboard rewards, students who do not like competition, and students who perceive the leaderboard to be using performance-contingent rewards to control their behaviour.

Even though the theory surrounding leaderboards causes concern, a range of literature (e.g., Aldemir et al., 2018; Cheong et al., 2013; De Schutter & Abeele, 2014) shows that students generally consider leaderboards to be motivating, enjoyable, and engaging. The cognitive evaluation theory (Ryan & Deci, 2009) aspect of SDT suggests that the rewards used by a leaderboard positively impacts intrinsic motivation if they are perceived as informational; informational rewards allow people to feel in control and competent. The literature (e.g., Çakıroglu et al., 2017; Gåsland, 2011; Mekler et al., 2013) provides support for leaderboards by arguing that they increase intrinsic motivation by fostering competence through achievement, relatedness through the shared community (O'Donnell et al., 2013; Sheldon, 2011), and autonomy by allowing students to self-assess their performance (Aldemir et al., 2018). Mekler et al. (2013) found that game elements such as points and leaderboards do not affect perceived autonomy, competence, or intrinsic motivation, but act as progress indicators that guide and enhance the participants' experience.

Richter et al. (2015) explains that "combining a leaderboard with points adds a social dimension with an unknown effect on motivation: it may either promote intrinsic motivation by experiencing competence, or reduce intrinsic motivation, if perceived as controlling" (p. 37). The LLOS, presented in Section 2.2.2, appears to be a suitable instrument to examine how leaderboards affect FL motivation. The LLOS measures three types of extrinsic motivation and three types of intrinsic motivation. As leaderboards use external rewards such as points and rank to control behaviour, the leaderboard appears to align with the least autonomous form of extrinsic motivation, external regulation. However, if the leaderboard rewards are perceived as informational, the leaderboard could align with the intrinsic motivation accomplishment subscale as it reflects achieving L2/FL goals.

2.4.3 Emotions and Attitudes towards Leaderboards

The literature examining leaderboards is clearly mixed and further research is required. Students' emotional reaction to leaderboards could provide insight into how

leaderboards affect performance and motivation. Koster (2005) explains that "a game is a system in which players engage in an abstract challenge, defined by rules, interactivity, and feedback, that results in a quantifiable outcome often eliciting an emotional reaction" (p. 34). In an educational context, it is important to understand emotions as they can affect learning, performance, motivation and personal growth (Heckhausen, 1991; Zeidner, 1998); some emotions could suggest engagement, other emotions could suggest disengagement (Pekrun, Frenzel, Goetz, & Perry, 2007).

According to Izard (1991), understanding the outcome effect of emotions is challenging as emotions do not occur in isolation, and they affect people differently. The control-value theory of achievement emotions (Pekrun et al., 2007) provides guidance to understand the relationship between an emotion and its outcome effect. The theory defines achievement emotions "as emotions tied directly to achievement activities or achievement outcomes" (Pekrun et al., 2007, p. 15). However, not all the emotions in an educational context are achievement emotions; some are social emotions which overlap with achievement emotions; the control-value theory of achievement emotions takes it into account.

Pekrun et al. (2007) explain that there are three important dimensions of achievement emotions: object focus, valence, and activation direction. The object focus refers to the origin of the emotions as either being activity-related or outcome-related. Activity-related achievement emotions are experienced while doing the learning activity. Outcome-related achievement emotions are experienced when academic goals are met or not met. Next, emotions are classified based on their valence of being either positive or negative. Finally, the activation direction refers to whether an emotion leads to activation or deactivation. Positive emotions do not necessarily lead to activation, and negative emotions do not necessarily lead to deactivation (Saldaña, 2009). For example, the positive emotion excitement could lead to activation, whereas the positive emotion pride could lead to deactivation. The full control-value theory also takes into account that emotions can be determined based on factors such as whether the student feels in control of an activity, whether the activities are important to them, genetic and physiologically disposition, socio-historic context, and that emotions can vary in intensity.

Gamification implementations result in a variety of emotions depending on feelings of perceived success or failure (Dominguez et al., 2013). Studies (e.g., Cheong et al., 2013; Philpott, 2015a) show that students at the top of a leaderboard feel more positive emotions towards leaderboards, whereas students at the bottom feel more negative emotions. In contrast, Bursztyn and Jensen (2015) found that high school students felt embarrassed by appearing at the top of a class leaderboard and reduced their performance to avoid appearing on the leaderboard. Negative feelings associated with the competitive and the comparative nature of leaderboards in an educational setting are a reoccurring theme in the literature (e.g., Barata et al., 2013; Charles et al., 2011; Domínguez et al., 2013; Nicholson, 2013). Aldemir et al. (2018) found short deadlines of 2-3 days led to feelings of fear and distress; however, the effect these emotions had on performance and motivation was not directly explored.

The use of leaderboards in a gamified course will have a a psychological effect on students (Cheong et al., 2013; Philpott, 2015a) with different students affected differently (Dubravac, 2012; Domínguez et al., 2013; Jia, Liu, Yu, & Voida, 2017; Wells & Skowronski, 2012) based on variables such leaderboank rank (Aldemir et al., 2018), social status (Bursztyn & Jensen, 2015), personality (Codish & Ravid, 2013) and motivational inclination (Heeter, Lee, Medler, & Magerko, 2011; Hew, Huang, Chu, & Chiu, 2016; Tan & Hew, 2016). Philpott (2015a) determined a range of emotions students feel when they looked at a leaderboard in an EFL course at a Japanese university. The emotions were proud, satisfied, happy, sad, disappointment, embarrassment, ashamed, regret, nothing special. Based on Philpott (2015a), this thesis aims to determine if the emotions derived from a leaderboard in a gamified EFL course lead to activated or deactivated behaviour, and whether the emotions provide insight into the effect leaderboards have on performance and L2/FL motivation.

2.5 Quests in Gamified Instructional Design

Quests are a gamification component that strive to challenge, engage, and entertain a player (Kapp, 2012). On the surface, quests are thematically-packaged challenges that include an objective, instructions, and a corresponding reward that is given to the

player if they can complete the quest successfully (Ashmore & Nitsche, 2007; Dubravac, 2012). Below the surface, the purpose of the quest could be to encourage the player to learn or practice a new skill or have some type of meaningful experience. When utilised in an EFL context, on the surface, quests appear as language learning activities; below the surface, quests have the potential to foster intrinsic LF/FL motivation if designed appropriately (Sheldon, 2012).

The use of quests in an educational setting is referred to as quest-based learning (QBL) (Haskell, 2012). QBL can be a stand-alone pedagogical approach, or it can be part of a gamified class, interconnected with other gamification components such as leaderboards, points, levels, badges, and a narrative (Sheldon, 2011). Teachers generally prepare a number of quests that strive to provide relevant learning experiences for their students who complete them over a semester or an academic year (Sheldon, 2011). Sullivan, Mateas and Wardrip-Fruin (2009) outline two types of quest-design structure: task-based and goal-based. Task-based quests clearly describe a list of tasks that must be completed for the quest to be completed, whereas goal-based quests present the objective of the quest and allow the player the freedom to decide how they complete it. The underlying assumption about QBL is that, because the quests appear game-like, this increases student engagement and intrinsic motivation as young people are more motivated by games than traditional learning content (Sheldon, 2012).

Since Sheldon (2012) clearly outlined what quests are, and claimed that they are an intrinsically motivating pedagogical approach, interest surrounding their utility has increased. Unfortunately, research supporting their utility has not proceeded at a rate commensurate to their adoption. There is currently a lack of literature to explain how quests should be designed to support FL acquisition. There is also a lack of literature to strongly support Sheldon's claims about quests being able to foster intrinsic motivation for any educational context, especially EFL. The psychological and EFL pedagogical literature (e.g., Krashen, 2009; Ryan & Deci, 2009; Seligman & Csikszentmihalyi, 2000; Swain, 1985; Willis & Willis, 2007) that surrounds QBL suggests that QBL can achieve pedagogical goals and intrinsically motivate students if the quests foster suitable levels of competence, autonomy, and relatedness. The following sections explain the surrounding literature, detail issues pertinent to the

Japanese context, present the QBL research that has occurred so far, and provide a rationale for further research to occur.

2.5.1 Quest-based Learning as an Approach for EFL Pedagogy

Literature that specifically explains how quests should be designed for a FL course is lacking; however, TBLT and L2 acquisition theories provide guidance. Shintani (2011) explains that TBLT is a communicative approach to teaching a foreign language that borrows from a range of L2 acquisition theories such as the Interaction Hypothesis (Long, 1996), the Cognitive theory of L2 learning (Skehan, 1998), and the Cognition Hypothesis (Robinson, 2003). TBLT provides a broad framework for designing quests that are pedagogically effective. Proponents (e.g., Willis & Willis, 2007) of TBLT say that engaging students in tasks that allow real-life, authentic language use is an effective way to foster L2/FL development while motivating students. Shehadeh and Coombe (2012) describe a task as an activity that has a non-linguistic goal, with a clear outcome, which requires any or all four of the language skills to be accomplished in a way that reflects real-world language use. Quests are similar to tasks as they are both encapsulated learning activities. The main difference between a quest and a task is that a quest generally has some gamification components attached such as a narrative, theme, or points. Findings from the large body of TBLT literature (e.g., Breen, 1987; Samuda & Bygate, 2008; Van den Branden, Bygate, & Norris, 2009; Van Avermaet & Gysen, 2006) that has accrued since the late 1980s provide guidance for designing quests that are pedagogically effective.

The TBLT literature (e.g., Ellis, 2003; Nunan, 1989) states that the context, culture, sequencing, and grading of tasks requires consideration; these variables are also important for QBL (Hamari et al., 2014). Suitably challenging quests need to be designed for the specifics of the context. Quests designed for high school students are probably not suitable for university-level students, and at university-level, questing may be suitable for some courses, but not others. Also, within a class, there will be students of different genders, personalities, nationalities, and intelligences which could respond differently to different aspects of questing (Koster, 2005). How

tasks are sequenced is important, they should be presented and completed in a way that is conducive to learning the learning context; levels can be used to control the flow and progression of quests. Finally, the grading of tasks should provide meaningful feedback to the students (Ellis, 2003). Kapp (2012) explains that feedback in gamification turns the game into a learning experience. Students need to be able to receive meaningful feedback about their completed quests.

From the SLA field, Krashen (2009) and Swain (1985) provide guidance for what type of activities should be incorporated into a quest for L2/FL acquisition to occur. Krashen's Input Hypothesis suggests that L2 learners should be exposed to large amounts of comprehensible input through listening and reading in the target language. Swain's (1985) Comprehensible Output Hypothesis states that L2 acquisition occurs when learners are outputting language. Swain (1995) defined three functions of output: (1) the noticing/ triggering function, (2) the hypothesis-testing function, and (3) the metalinguistic function. Simply stated, output enables learners to identify gaps between what they want to say and what they can say. Once a gap has been identified, learners then attempt to bridge the gap by applying a learning strategy such as using a dictionary or asking a teacher for help (Swain, 2000). Even though the importance of output is debated in the literature (e.g., Krashen, 2003), it appears to be a logical aspect of foreign language acquisition that can work alongside Krashen's hypotheses (Liu, 2015). TBLT literature alongside the Input Hypothesis and the Output Hypothesis provide a theoretically supported framework for designing quests that support L2/FL acquisition.

2.5.2 Fostering Intrinsic Motivation using Quest-Based Learning

The literature that surrounds QBL suggests that Sheldon's (2012) claim that QBL is a motivating pedagogical approach can be true in the EFL context if the design of the quests incorporates sound academic theory. This section uses the three tenets SDT as a framework to explain how QBL can foster intrinsic motivation. Aspects of positive psychology, the L2MSS, and the international construct are then introduced to show how they can strength the SDT framework. When possible, education-based QBL literature is detailed to support the argument that QBL can be intrinsically

motivating. The lack of literature shows that further exploratory research is required, especially in the EFL context.

Self-determination theory (Ryan & Deci, 2009) asserts that activities that are suitably challenging are intrinsically motivating. Lambert, Gong and Harrison (2015) support the notion that quests that are suitably challenging foster intrinsic motivation. They compared two groups of participants in an educational technology course: the treatment group was subjected to QBL, and the control group completed the course in the traditional manner. The Intrinsic Motivation Inventory (Ryan, 1982) was administered to both groups of participants as part of the quasi-experimental design to determine if QBL was more motivating than a traditional course. Lambert et al. showed that students generally viewed QBL as more valuable and useful for learning compared to a traditional approach in an educational technology course. The experimental group reported a higher score on the factor that measured enjoyment/competence, and statistically significant higher scores for the factors that measured value/usefulness, and effort. The researchers hypothesised that one of the reasons for the high results was because the participants had to demonstrate more competence in order to successfully complete the quest compared to a traditional course in which the participants would just receive their mark and then move on, and because of the autonomous nature of the quests.

Çakıroglu et al. (2017) also support the notion that quests that are suitably challenging foster motivation. They found positive student perceptions towards the use of quests as a pedagogical approach in an undergraduate information and communications technology course. Many of the participants explained that the quests required them to deeply consider how they were going to complete the quests. This not only made them feel motivated while completing the quests but it also "facilitated deep learning" (p. 103). The participants explanation about how the feeling of motivation stemmed from requiring deep consideration suggests that the quests were motivating because they were suitably challenging. The findings from Lambert et al. (2015) and Çakıroglu et al. (2017) alongside SDT theory (Ryan & Deci, 2009) suggest that quests that are suitably challenging will be intrinsically motivating.

Self-determination theory (Ryan & Deci, 2009) asserts that activities that support autonomy are intrinsically motivating. QBL is suitably positioned to foster the autonomy tenet of SDT in two obvious ways: first, by allowing the students a choice of what quest to work on; second, by allowing students the autonomy to make meaningful choices about how they complete the quests they work on. The type of autonomy quest choice can provide is strongly encouraged in the L2/FL literature (e.g., Benson, 2007; Pemberton, Toogood, & Barfield, 2009) as an approach to foster resilient long-term L2/FL motivation. Benson (2011) explains that autonomous learning behaviour can be developed by allowing and encouraging students to make certain choices about the learning content used as part of the course. Game designer McGonigal (2011) says that the successful completion of a meaningful quest that provided autonomy and required a suitable degree of competence should reward the player enough intrinsically so that extrinsic rewards such as points or badges are inconsequential.

Providing a selection of quests to complete is a common aspect of QBL. Haskell (2012) conducted an experiment to determine the attractive and interesting features of quests that led to quests being chosen, completed, and highly rated by students in a university-level introductory educational technology course. 66 quests, divided into levels, covering a range of appropriate topics were developed and delivered to students on the QBL website Rezzly (http://rezzly.com/). The attractiveness was determined using a formula that calculated variables related to a quests ability to capture one's interest, sustain one's effort, and provide a personally relevant learning experience to the student. The results of Haskell's study show that the students' initial perception of the quest was important; the inclusion of various multimedia such as blogging, podcasting, and videos appeals to students; students are attracted to quests that look like they can be completed quickly, and to task-based quests that clearly show how they should be completed. Haskell found that, even though task-based quests are initially more appealing, the students enjoyed completing goal-based quests more, possibly due to them fostering stronger feelings of autonomy and competence.

De Schutter and Abeele (2014) measured students' attitudes towards 12 components incorporated into their gamified undergraduate liberal education course. For each

component, students assigned a score for different aspects related to motivation, enjoyment, engagement, and should the component be removed from the course. The mean scores for the component about allowing a choice of quests to complete was noticeably higher than all the other components. The students strongly agreed that choosing quests was motivating, enjoyable, and engaging. The students disagreed that the being able to choose which quest to work on aspect should be removed from the course. The literature (e.g., Benson, 2011; Haskell, 2012; McGonigal, 2011; De Schutter & Abeele, 2014; Seligman & Csikszentmihalyi, 2000) about quest choice suggests that by allowing students a choice of what quest to complete, and the autonomy to decide how it will be completed, intrinsic motivation can foster; however, this is not always true for all students. Çakıroglu et al. (2017) showed a small number of students perceived QBL to be boring due to the large number quests.

Literature that examines the use of QBL in a Japanese EFL context does not exist. However, surrounding literature can provide guidance. Philpott (2015b) conducted a mixed methods study to explore student opinions towards self-access learning tasks in an EFL course at a Japanese university. The self-access tasks are similar to quests without the surrounding gamification elements such as a narrative and XP. Philpott prepared 20 tasks and grouped them into six thematically similar categories which acted as levels. Students were told to complete one task a week, and they could progress to the next level when they had completed at least two tasks in a level. At the end of the course, the students were asked to write any opinions they had towards the tasks. The results showed generally positive comments towards the tasks. Different students enjoyed different types of tasks. Interestingly, half of the students said they liked the video making tasks and half of the students said they did not like the video making tasks.

Philpott's findings support a reoccurring theme in the literature (e.g., Lambert, 2017; Perry, 2015) about gamification components, in this case QBL, affecting different students differently. Through a series of studies (e.g., Lambert & Ennis, 2014; Lambert, Gong, & Harrison, 2015; Lambert, 2017) that measured individuals subjective experiences towards QBL, Lambert (2017) determined that extrinsically motivated students were less motivated by the class that incorporated QBL compared

to their extrinsically motivated counterparts in the traditional control class. Lambert also found that the students who already possess autonomous learning behaviour enjoyed QBL more than the students who do not possess autonomous learning behaviour. The students who did not already possess autonomous learning behaviour found QBL to be more challenging. Lambert's findings in conjunction with SDT suggest that intrinsically motivated students will respond positively towards QBL, whereas the extrinsically motivated students will not respond positively.

Self-determination theory (Ryan & Deci, 2009) asserts that activities foster relatedness are intrinsically motivating. As collaboration and cooperation have been hallmarks of many games, QBL appears suitably positioned to foster the relatedness tenet of SDT through collaborative and cooperative quests. Literature directly supporting this claim is lacking; however surrounding literature provides support. Lambert et al. (2015) shows that students in an Educational Technology course who were subjected to QBL reported higher levels of relatedness compared to the students in the traditional version of the course, which resulted in a statistically higher effort which also corresponded with higher levels of enjoyment. In gamified contexts, players generally prefer collaborating on activities rather than doing them individually (Lounis, Pramatari, & Theotokis, 2014). The preference for collaboration and cooperation is beneficial as it can increase performance and motivation (McGonigal, 2011; Ryan, Rigby, & Przybylski, 2006; Sheldon, 2011). In education, cooperation has been linked to increased academic achievement (Ames & Felker, 1979) and heightened self-esteem (Myers & Lamm, 1976). Self-esteem is a sign of psychological well-being and SDT asserts that psychological well-being leads to intrinsic motivation (Ryan & Deci, 2009). The QBL literature about collaboration is mixed but lacking.

Gamified classrooms have commonly encouraged collaboration by grouping students into teams, also known as guilds. Guilds encourage collaboration through teamwork in which students work together to achieve goals. The literature (e.g., Sheldon, 2012) about the use of guilds in educational gamification is mixed but provides warnings for teachers looking to utilize guilds in their gamified class. De Schutter and Abeele (2014) measured student attitudes towards 12 components incorporated into a gamified undergraduate liberal education course. The use of guilds was the third

lowest ranked component; however, the mean scores showed that students slightly agreed that guilds were a good component of the course because they resulted in feelings of motivation, enjoyment, and enjoyment. The students' attitudes towards whether guilds should be a part of the course were mixed.

The literature (e.g., Aldemir et al., 2018; De Schutter & Abeele, 2014; Kapp, 2012) details some aspects of guilds that can negatively affect intrinsic motivation and learning. First, guilds can allow social loafing to occur; a situation where students reduce the amount of work they do because they know their group members will pick up their slack (Hoon & Tan, 2008; Kapp, 2012). Second, some students prefer collaboration over competition but disapprove of working in teams as low performing students decrease the team's chance of success (Aldemir et al., 2018). SDT suggests that students' intrinsic motivation could be negatively impacted if they are forced to work with people who they may not want to work with by reducing their feeling of autonomy, not allowing the relatedness tenet to foster as meaningful interaction might not occur, and failing to foster a sense of competence if students relinquish their responsibilities to other students. The literature (e.g., Aldemir et al., 2018; De Schutter & Abeele, 2014; Sheldon, 2012) alongside SDT suggests that guilds may not be an effective approach to foster meaningful collaboration among university level students.

Collaboration and relatedness can be fostered through the design of the QBL system and through individual quest design. SDT suggests that collaboration should be fostered in an autonomous manner. Students should not constantly be forced to work with people who they may not want to work with; however, they should be provided extensive opportunity to collaborate. In an EFL context, environments that foster collaboration can reduce communication-related anxiety, which leads to increased self-confidence and motivation (Koga, 2010). The design of the quests should entice students to want to collaborate; however, students who do not want to collaborate should be able to choose a different quest or be able to complete the quest in a manner that does not require collaboration. Ideally, once a student has successfully completed a collaborative quest or seen the enjoyment another student received from completing a collaborative quest, they will feel more motivated to complete collaborative quests in the future (McGonigal, 2011). It appears that quest can

increase students' L2/FL motivation through quests that foster the relatedness tenet of SDT; however further research is required to determine if the assumption is true.

Literature from the positive psychology research field provides further guidance for designing quests that foster intrinsic motivation. First written about by Martin Seligman (1998) in opposition to the pathologically focused psycho-analysis and manipulative behaviourism, positive psychology proposes that psychologists should focus on what makes people happy and fulfilled. Positive psychology is defined as "the scientific study of positive human functioning and flourishing on multiple levels that include the biological, personal, relational, institutional, cultural, and global dimensions of life" (Seligman, 2013, p. 2). One of its fundamental assertions is that it is just as important to help and support the lives of healthy people as is helping those in need (Seligman & Csikszentmihalyi, 2000).

Seligman (2008) explains that there a various types of happy lives people can live, but it is those lives that are made of meaningful and engaging activities which lead to the most resilient happiness and well-being. Psychological well-being does not have to happen by chance. It can be fostered by doing activities which provide intrinsic rewards while not valuing extrinsic rewards so highly. When people do activities that generate intrinsic rewards such as positive emotions, personal strength, and social connections (Lounis et al., 2014), the enjoyment they get is enough to make humans truly happy. Hundreds of studies and experiments show a link between hard work, happiness, and intrinsic reward (explained by McGonigal, 2011).

Positive psychology can be built into quests to help foster intrinsic motivation. Seligman's (2008) PERMA framework outlines five aspects to incorporate: positive emotions, engagement, relationships, meaning, and accomplishment. The PERMA framework suggests two additional approaches to fostering intrinsic motivation that are not specifically addressed by SDT. First, quests should be interesting and enjoyable. Reeve (1989) found that "interest contributes to intrinsic motivation by arousing the initiation and direction of attention and exploratory behavior, while enjoyment contributes to intrinsic motivation by sustaining the willingness to continue and persist in the activity" (p. 83). Second, quests should be meaningful. Lyubomirsky (2007) recommends that quests should have meaningful goals and

facilitate meaningful experiences on the way to completing the goal.

Dörnyei's (2005) L2 Motivational Self System (L2MSS) also suggests an approach to increasing intrinsic motivation that is not specifically targeted by SDT and positive psychology. The L2MSS evolved from the social-educational model to provide an updated theory on L2 motivation. The L2MSS is based on Higgins's (1987) self-discrepancy theory and Markus and Nurius's (1986) concept of 'possible selves' and includes three components that are not mutually exclusive: ideal L2 self, ought-to L2 self, and L2 learning experience. When language learners imagine the type of language speaker they want to become in the future, this person is their ideal L2 self; learners are motivated to become this person. It has been shown to be an important indicator of intrinsic motivation (Ryan, 2009). The L2MSSS has overlaps with SDT; however, the ideal L2 self construct provides another avenue for fostering intrinsic motivation. To further support the ideal L2 self construct, Yashima's (2002, 2009) international posture construct recommends that activities should increase students' interest in the English-speaking world outside of Japan.

The effective implementation of QBL in a Japanese EFL context requires special pedagogical and psychological considerations. Burrows (2008) in relation to TBLT in a Japanese university context outlines some important factors to consider: learning styles, learning expectations, socio-cultural differences, and the structure of TBLT. Confucius values instilled in Japanese students have influenced them to respect, obey, and rely on instruction from their teachers (Stapleton, 1995). Japanese students enter university after completing high school which is based on teacher-centred learning. QBL focuses heavily on student-centred learning and students may not understand it or do not feel comfortable with it to begin with (Fisher, Hafner, & Young, 2007). Therefore, students who are accustomed to a teacher-centred education environment may need psychological deconditioning (Holec, 1985) before they can succeed in an autonomously natured QBL environment. Also, as Japan is a collectivist country with many Japanese people feeling shy in social situations (Zimbardo, Pilkonis, & Marnell, 1977), the collaborative and competitive social aspects of a gamified class may lead to feelings of uncomfortableness. Even though the students may feel uncomfortable, the structure of QBL can allow for the scaffolding and the tailored introduction of activities which transition students from a

teacher-centred learning environment to a more autonomously natured student-centred learning environment.

2.5.3 Opinions and Perceptions of Quest-Based Learning

Student opinions and perceptions towards QBL could provide insight into what aspects of QBL impact motivation and learning. Opinions refer to what students think about specific aspects of QBL, whereas their perceptions refer to how they regard QBL. Student perceptions towards QBL could be affected by their personal history, future goals, biological makeup, course expectations, culture, and the opinions of people they consider important according to Nelson (2008). Unfortunately, few studies explore students' opinions and perceptions towards specific gamification components such as QBL (Rapp, 2015). Some studies (e.g., Aldemir et al., 2018; Buckley et al., 2017) have addressed the gap in the literature; however, not only did they did not analyse QBL, they did not use an EFL context for the studies. Section 2.5.2 showed that the literature (e.g., Çakıroglu et al., 2017; Haskell, 2012; Lambert et al., 2015; Lambert, 2017; Philpott, 2015b; Sheldon, 2012) that did explore students' opinions and perceptions was able to provide rich insight into various aspects of QBL. Both Lambert et al. (2015) and Çakıroglu et al. (2017) successfully used psychometric questionnaires to collect data; however, Çakıroglu et al. led to richer findings by also collecting data through interviews.

This study intended to address the gap in the literature about QBL in an EFL context with a focus on its viability to be an effective approach to pedagogy as a pedagogical approach, and its effect on intrinsic FL motivation. Qualitative data were collected and analysed in an exploratory manner that allowed the participants' opinions and perceptions to naturally surface. The participants shared their opinions and perceptions towards individual quests, QBL as a pedagogical approach, quest choice, and collaboration. Due to the rich findings from Lambert et al. (2015) and Çakıroglu et al. (2017), the LLOS (see Section 2.2.2) was used as a psychometric tool to measure motivation, and semi-structured interviews were used to further explore and triangulate findings.

2.6 Theoretical Framework

Gamification implementations need to be designed based on the specifics of the context (Buckley et al., 2017; Garland, 2015). There are various academic theories and frameworks that can be used to guide the design of this implementation that occurs in an EFL course at a Japanese university. Richter et al. (2015) suggest that "a conceptual consolidation of theories may aid to carefully craft reward and incentive mechanisms to increase short-term and long-term performance and promote game persistence" (p. 37). The theoretical framework for this study, presented in Figure 2.1, shows how a gamification framework, psychological theories, and SLA theories are combined to construct the gamified course. The theoretical framework then shows the five data collection instruments, the variables of interest, and what will be measured.

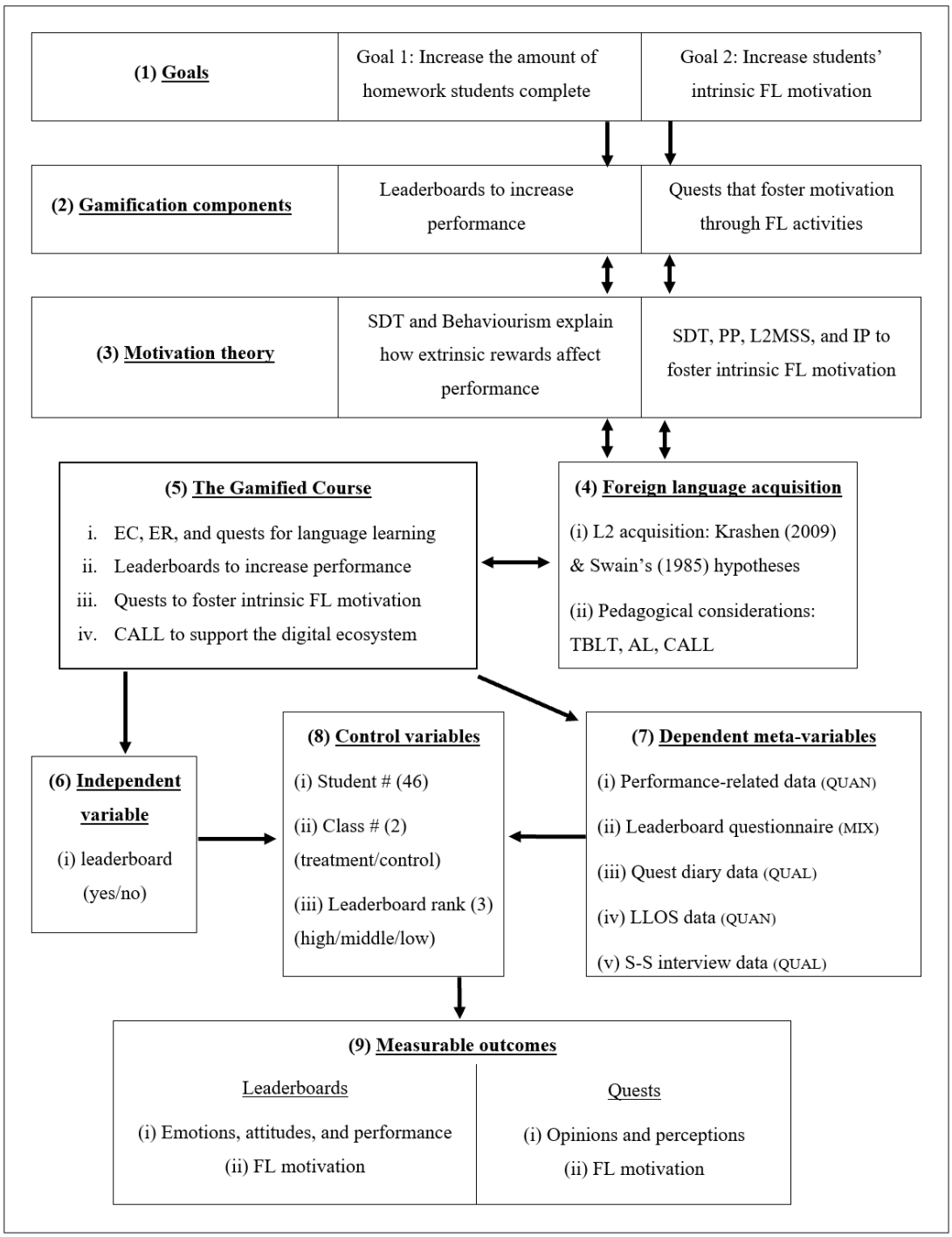


Figure 2.1. Theoretical framework diagram.
Note: The theoretical framework provides an overview of the important aspects of the study.

Box (1) represents the first step of designing a gamification implementation, defining the goals. The goals for this implementation stemmed from the researcher's desires to increase students' short-term performance and long-term intrinsic FL motivation. Box (2) shows the main gamification components that were incorporated into the gamified course. Werbach and Hunter's (2012) framework, explained in Section

2.3.1, guided the choice of components. Leaderboards were included to increase performance through competition, feedback and rewards. Quests were included due to their potential to increase intrinsic FL motivation through language learning activities.

Box (3) shows the motivation theories that explain how the gamification components will achieve their goals. For the leaderboard, SDT suggests that the participants will either be driven by the extrinsic rewards of the leaderboard, such as points and rank, or the leaderboard will inadvertingly encourage performance by supporting intrinsic motivation. Behavourism suggests that the more likely outcome is that the students' performance will increase either to get the extrinsic rewards, or avoid being punished by having their negative results displayed on the leaderboard. Behaviourism suggests that even if the participants are motivated to get the extrinsic rewards, their performance will eventually decrease as they lose interest in the extrinsic rewards (Chen, 2003; Ferster & Skinner, 1957).

Self-determination theory was used as the central framework to design quests that support intrinsic motivation. The quests aimed to be suitably challenging, support autonomy through quest choice, and encourage relatedness through collaboration. Three other theories were also incorporated into the design of some of the quests to further attempt to increase intrinsic motivation. Positive psychology literature (e.g., Seligman, 2008) encouraged designing quests that would be enjoyable, interesting, and personally meaningful. The ideal L2 self construct was borrowed from the L2MSS (Dörnyei, 2005) and incorporated into some of the quests to encourage the students to think about their future, and contemplate what they need to do to achieve their future goals. Some of the quests were designed to support the international posture construct (Yashima, 2002; 2009) by encouraging the students to become interested in the world outside of Japan.

Box (4) shows the SLA theories that were incorporated into the design of the gamified course to ensure the course was suitable for FL acquisition. Krashen's (2009) input hypothesis (explained in Section 2.5.1), and reading hypothesis explain how L2 acquisition occurs. The reading hypothesis (Krashen, 2009) asserts that when students are reading in the target language without a struggle, being about to

understand at least 95% of the words in a text, relatively efficient language acquisition can occur. The literature (i.e., Krashen & Mason, 2015; Nation, 2014) supports the claim and estimates that for every hour of independent English reading a student does, their TOEIC score test increases by half a point. Swain's (1985) comprehensible output hypothesis, explained in Section 2.5.1, provides further support by asserting that L2/FL acquisition occurs when learners are outputting language.

Box (4) also shows the main pedagogical considerations for the gamified course. TBLT literature (e.g., Ellis, 2003) provided the framework to design quests that incorporate motivation theory while supporting FL acquisition. Autonomous learning literature (e.g., Benson, 2007) favoured the inclusion of learning materials that support autonomous learning and encourage the students to become autonomous learners. Findings from the CALL literature (explained in Section 2.3.2) encouraged the inclusion of learning activities that had large online components. The preference was due to the ability of CALL to efficiently deliver learning content, assess completed work, and provide feedback to the student and the teacher.

Box (5) shows the main elements of the gamified course. The arrows moving from Box (1) to Box (5) show how the gamified course is designed based on gamification, motivation, L2/FL acquisition theories and approaches. The bidirectional arrows from Box (2) to Box (5) reflect that all aspects support and not hinder each other. The three language learning activities are supported by the SLA literature. English Central (EC) and extensive reading (ER) both provide large amounts of suitable levelled input which aligns with Krashen's input and reading hypotheses respectively. The quests foster comprehensible output recommended by Swain (1985) as they require the students to use their English skills in an uncontrolled environment in which they will have to identify and close the gaps between what they want to say and what they can say before they can complete the quest. All three activities are CALL-based and promote autonomous learning. EC and ER are homework activities the students worked on in the previous semester. The continuation of the activities preserves the ecological validity of the study. Leaderboards are primarily used to encourage students to complete all the language learning activities, and quests are primarily used to foster intrinsic FL motivation. A

range of CALL-related websites and software were used to create a digital ecosystem that included the leaderboard and the QBL system.

Boxes (6), (7), and (8) show the important variables for this study. Box (6) shows the independent variable, a leaderboard, that will be applied to the treatment group to see how it affects student performance. Box (7) list the five dependant meta-variables that are analysed in this study; subsumed under the meta-variables are many other variables that target aspects of leaderboard and quests that of importance for this study. Box (8) shows the control variables that are studied in relation to the independent and dependant variables. The number in brackets represents the number of different entities for each variable. Box (9) shows the measurable outcomes that will be used to answer the research questions.

2.7 Summary

Second and foreign language motivation continues to be a frequently examined topic in the Applied Linguistics research field as it is considered a crucial factor in determining L2/FL success (Brown, 1981; Buckley & Doyle, 2014). Gardner and Lambert's (1959, 1972) seminal finding that a learner's attitude toward the L2 community significantly affects L2 learning behaviour set the tone for early L2 research. Research then moved from focusing on ethnolinguistic communities of language learners to more situated and specific learning contexts such as language classrooms. The cognitive-situated period of L2/FL motivation research applied mainstream psychological theories such as Deci and Ryan's (1985) SDT to the L2/FL context and determined that learning activities that foster competence, autonomy, and relatedness lead to intrinsic motivation towards (Hiromori, 2003). From around the year 2000, the cognitive-situated period of L2/FL research merged into the process-oriented period of research that explores how L2/FL motivation changes over time (Dörnyei & Ushioda, 2011); multiscale psychometric questionnaires such as the LLOS are instruments created for the purpose of measuring students' L2/FL motivation.

Foreign language motivation research shows that the Japanese peoples' struggle with

becoming confident English communicators is attributed to a lack of L2/FL intrinsic motivation (Ushioda, 2013). In an attempt to create learning experiences that motivate and engage Japanese students, this thesis proposes that gamification techniques be applied to the structure of an EFL course. The mainstream gamification literature (e.g., Werbach & Hunter, 2012) explains that gamification has been successfully used in a variety of situations to motivate behaviour. However, the other literature (e.g., Cook, 2018; McGonigal, 2011) warns that, if gamification replaces intrinsic motivation with extrinsic motivation, it could lead to a range of negative outcomes such as lower quality of work, or future demotivation. Literature supporting or refuting the efficacy of gamification in an educational context is far from conclusive; it often fails to isolate specific gamification components, rarely occurs in an EFL context, and has not occurred in the Japanese context. This thesis aims to determine how the gamification components, leaderboards and quests affect Japanese students' performance-related behaviour and their L2/FL motivation in a gamified EFL course.

Leaderboards are frequently used in gamification implementations. If the assertion that gamification replaces intrinsic motivation with extrinsic motivation is true, leaderboards could be a central cause. Behaviourism and SDT explain how leaderboards use extrinsic rewards such as points and rank to influence behaviour. The literature (e.g., Deci et al., 1999; Kapp, 2012; Ryan & Deci, 2009) suggests that, if students feel that the extrinsic rewards provided by a leaderboard are controlling their behaviour, rather than supporting their behaviour, this negatively impacts their L2/FL motivation. On the other hand, a range of literature (e.g., Çakıroglu et al., 2017; Mekler et al., 2013) argues that leaderboards foster L2/FL intrinsic motivation by fostering competence through achievement, relatedness through the shared community, and autonomy by allowing students to self-assess their performance.

Students generally like the use of leaderboards in their academic courses, and leaderboards often successfully result in a range of performance increases that the teacher may have been targeting. However, this is not true for all students (e.g., Aldemir et al., 2018; Tan & Hew, 2016). Low performing students are more likely to not respond positively to leaderboards, compared to the higher performing students. There is also concern that leaderboards do not represent learning and they limit

performance by creating artificial goals. Literature supporting any of the positive or negative assertions about leaderboards in an EFL context is extremely lacking. It is important for teachers to understand the possible positive and negative ramifications of using leaderboards in educational contexts. Therefore, this thesis aims to determine how leaderboards affect students' performance and L2/FL motivation when applied to a gamified EFL course.

Quests have been frequently used in traditional games for a long time to engage players; however, the use of quests in gamified learning contexts to engage students is a relatively new concept. Research supporting QBL in an EFL context is lacking; however, Sheldon's (2012) initial claim that students are intrinsically motivated to complete learning activities that are designed as quests is lightly supported in surrounding literature. A range of mainstream psychological, L2/FL learning, and gamification literature (e.g., De Schutter & Abeele, 2014; Ryan & Deci, 2009; Seligman, 2008; Willis & Willis, 2007) suggests that quests can foster the three tenets of SDT in various ways; for example, autonomy through quest choice, relatedness through collaboration, and competence by being suitably challenging.

As QBL literature is limited, students' perceptions of their QBL experiences allow important issues to organically surface. Studies (e.g., Çakıroglu et al., 2017; Haskell, 2012; Lambert, 2017; Lambert, et al., 2015; Philpott, 2015b; Sheldon, 2012) show that students generally have positive perceptions of QBL due to it promoting learner autonomy and interpersonal relationships. However, Lambert (2017) found the presence of a small number of students who did not have positive perceptions of QBL due to them not being comfortable with the autonomous nature of QBL. Lambert's finding could be important to consider for the Japanese context as students enter university accustomed to teacher-focused learning environments that are prevalent in Japanese high schools (Taguchi, 2005). Some Japanese students may not be comfortable with the autonomous nature of QBL. The QBL literature suggests that QBL has great potential to be an effective approach for EFL pedagogy, or at least, provide insight into aspects of L2/FL motivation. Therefore, this thesis aims to explore students' perceptions towards QBL, and determine the impact QBL has on students' FL motivation to determine its viability as an approach to EFL pedagogy.

The theoretical framework presented in Section 2.6 connects all the major aspects of the thesis. The framework is divided into two parts: the design of the gamified course, and the basic design of the study. The theoretical framework shows how gamification components, motivational theories, and SLA theories were combined to construct the gamified course used in this study. The theoretical framework then shows how five data collection instruments will be used to answer the research questions about leaderboards and quests.

Chapter 3. Methodology

3.1 Overview

This chapter details the quasi-experimental mixed methods research design that was used in this study to examine leaderboards and quests in an EFL course at a Japanese university. A convenience sample (suggested in Cohen, Manion, & Morrison, 2011) recruited participants (N = 46) from two EFL courses the researcher was already teaching at the host university. Data were collected over a 14-week period from two intact classes: Class 1 (n = 26) was the treatment group, Class 2 (n = 20) was the control group. The same pedagogical approach was used in both classes, including the use of quests; however, leaderboards were only used in Class 1. This chapter explains the three main components incorporated into the design of the gamified course: the gamification techniques, the homework, and the ecosystem that connected everything together. It then details the data collection instruments, the data collection procedures, and the data analysis methods that were conducted on each data set. It presents the ethical considerations surrounding this study and concludes with a chapter summary.

3.2 Research design

The research paradigm for this study is an original pragmatic approach that combines elements of postpositivism and social constructivism to develop a thorough understanding of the phenomena. Creswell (2014) explains that postpositivism is a deterministic philosophy that aims to identify the causes that affect an outcome through a suitable balance of quantitative and qualitative enquiry. Social constructivist theory encourages participants to interact with each other to develop an understanding of a subjective phenomenon. *Postpositivistic* research starts with a theory, whereas social constructivistic research inductively develops a theory based on data collected in the field. Pragmatism focuses on actions, situations, and consequences and it is the philosophical foundation for mixed methods research (Morgan, 2007; Tashakkori & Teddie, 2010). When an issue is embedded in a complex educational context, mixed methods research is particularly valuable as it can yield rich data that illustrates, clarifies, or elaborates on certain points (Mackey

& Gass, 2005; Mertens, 2009). The complexity of L2/FL motivation, the social nature of the leaderboards and quests, and the exploratory aspect of gamification research in an EFL context provide the rationale for a pragmatic mixed-methods research design.

This study employs a quasi-experimental mixed methods research design. It is quasi-experimental as random assignment of participants was not possible because the two participant groups were already members of intact classes that could not be separated (Cohen, Manion, & Morrison, 2007). The mixed methods design is hybrid in nature as both convergent and explanatory sequential approaches are used (Creswell, 2015). The convergent mixed methods design merges quantitative and qualitative data to answer the research questions in a way that is more reliable than if only one data type were analysed. The explanatory sequential aspect of the design collects qualitative data to further investigate or support initial findings by allowing triangulation to occur. Figure 3.1 outlines the research design process for this study. Stage 1 represents the data collection period. The data collection instruments collect either quantitative, qualitative, or mixed data. Stage 2 shows that after the data collection, each instrument is analysed in isolation. Stage 3 merges the relevant data from each instrument and then further explores or triangulates initial findings through semi-structured interviews. In Stage 4, the research questions are answered.

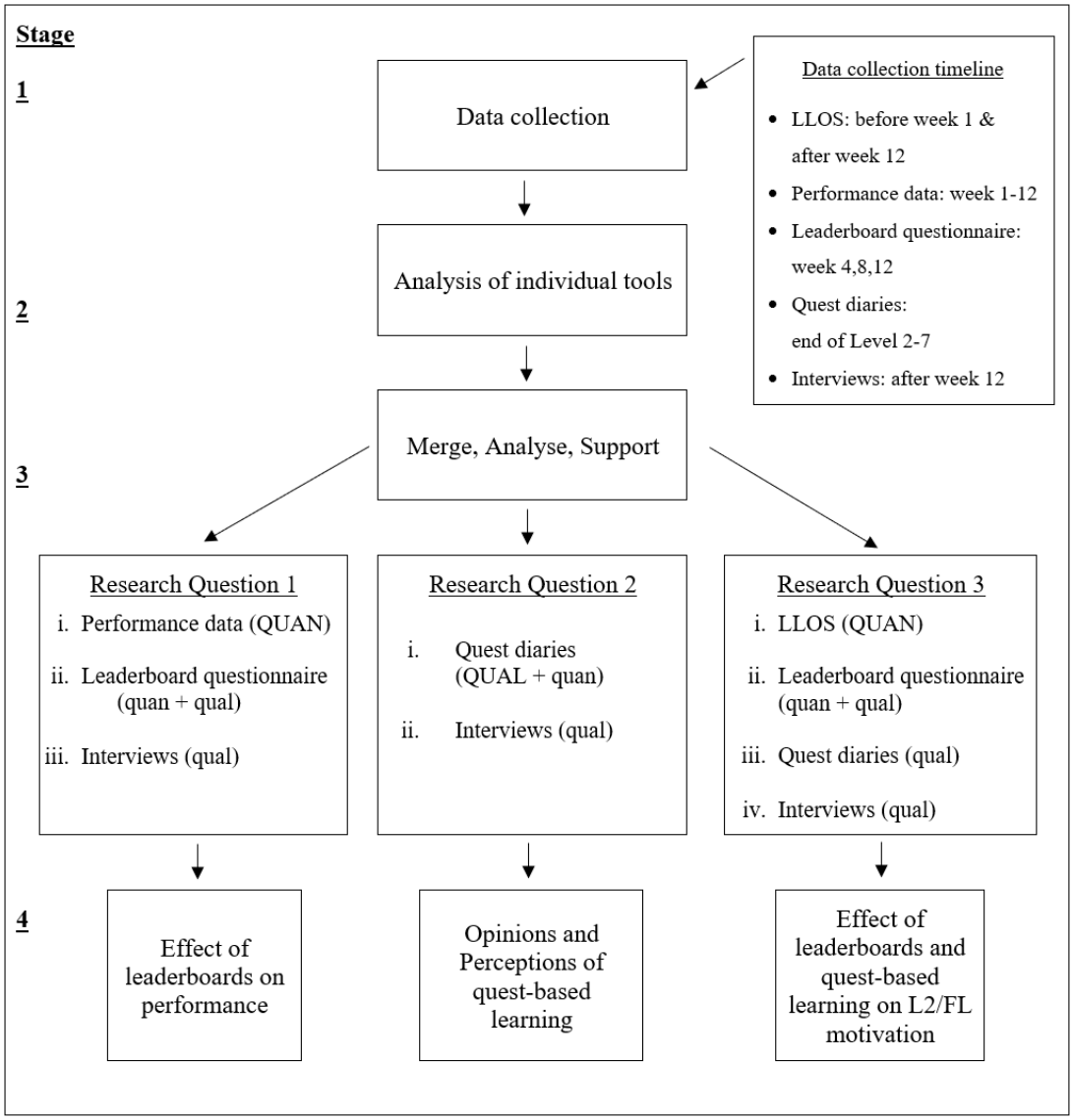


Figure 3.1. Research design process.
Note. 'QUAL' or 'qual' stands for qualitative research. 'QUAN' or 'quan' stands for quantitative research. Capital letters signify priority or increased weight. Lowercase letters signify lower priority or weight (Johnson & Christensen, 2004).

3.3 Participants

This study was conducted at a private, coeducational university located in the Kansai region of Japan. Two years of English education is mandatory for all undergraduate students at the university with each department administering a suitable English language program for their respective students. Referred to as *Gakubu English,* these classes occur twice a week, with one class taught by a Japanese teacher and the other class taught by a native speaker of English. However, students who can attain an acceptable TOEIC proficiency score are eligible to enrol in a 4-skills (reading,

writing, listening, and speaking) Intensive English course in the Language Center of the university. Students are streamed into Intensive English classes based on their TOEIC test results. There are two main reasons why students enrol in the Intensive English course: first, their course will be taught by only one teacher who is a native speaker of English; second, students can fulfil their English language requirements for graduation in one year, compared to *Gakubu English* that requires two years. The Intensive English classes are 90 minutes long, three-times a week (Monday-Wednesday-Friday), for two 14-week semesters. This study occurred in the second semester of an Intensive English course.

A convenience sample suggested in Cohen, Manion and Morrison (2011) was used to gather participants. The participants (N = 46) were all Japanese and came from two intact Intensive English classes. The researcher was the teacher of these Intensive English classes, and, as this study occurred in the second semester, the participants and the teacher had already established a relationship. The researcher considered the participants to be typical Japanese university students (explained in Section 2.5.2). Class 1 (n = 26) was made up of 12 males and 14 females, and Class 2 (n = 20) was made up of 8 males and 12 females. The classes met at different times of the day on the same day. For both classes, prior to the commencement of the first semester, the participants' TOEIC scores ranged between 575-625. The participants were 19 or 20 years old, second-year university students. For both classes, the participants came from a range of departments within the university such as business, economics, law, sociology, and humanities. Both classes had the same goals and assessment criteria. As these two classes had many similarities, they were suitable to use in a quasi-experimental study. Class 1 was chosen as the treatment group because it had more students. Informed consent was obtained from all the participants. The students who participated in this study will be referred to as participants from this point forward.

3.4 Materials

This section briefly explains the three main components incorporated into the design of the gamified course: the gamification techniques, the homework, and the ecosystem that connected everything. The researcher designed the gamified course to

achieve two goals: (1) motivate the participants to complete their homework; and (2) increase the participants intrinsic motivation towards studying English. The theoretical rationale supporting the design of the gamified course is found in Section 2.6.

The gamification techniques

The gamified course attempts to achieve goal (1) by using leaderboards to drive competition, and achieve goal (2) by encouraging the participants to complete specially-designed quests that foster L2/FL motivation. The leaderboard rankings were based on a point scoring system that was designed to encourage the participants to complete all aspects of their homework each week. The total weekly maximum score for each participant was 100 points, with each homework activity having a maximum score. The scores were used to create two ranked leaderboards, a weekly leaderboard, and an overall leaderboard. The weekly leaderboard showed each participant's score for each week, and the overall leaderboard showed each participant's accumulated points over the semester. The weekly maximum points were limited to 100 to avoid a situation in which an overwhelming point discrepancy between the high performing participants and the low performing participants developed. All participants were informed of the point scoring system; however, a leaderboard was only present in Class 1. Table 3.1 displays how a participant's weekly score was calculated based on the three homework activities: quests, extensive reading, and English Central. As part of the onboarding process, each participant in Class 1 received 100 points in Week 1. The maximum number of points each participant could achieve at the end of the study was 1300; 650 points for quests, 390 points for extensive reading using MReader, and 260 points for English Central. Table 3.2 provides a visual representation of the point accrual process.

Table 3.1

Weekly Points Scoring Table

Weekly points maximum score = 100 points (Quest + ER + EC)

Quest: 50 point maximum (due = before the start of Friday's class)

- 50 points for successfully completing 1 quest
- 30 points for 50% to 99% quest completion
- 15 points for 1% to 49% quest completion
- 0 points for no quest post on blog

Extensive Reading (ER) using MReader: 30 point maximum (due = Thursday 11:59pm)

- 30 points for reading 1 book and passing the quiz
- 15 points for attempting at least 1 quiz but not passing a quiz
- 0 points for not attempting quiz

English Central (EC) (100% = Watch 3 videos, Speak 3 videos): 20 point maximum (due = Friday 9:00am)

- 20 points for 100% completion
- 10 points for 50% - 99% completion
- 5 points for 1%- 49% completion
- 0 points for 0% completion

Note. The differing deadline times is due to that website having their own specific deadline times.

Table 3.2

Point Accrual Process

Week	Quests Points	Extensive Reading Max points	English Central Max points	Weekly Total Max Points
1	**LEVEL 1** 50 points	30	20	100
2 & 3	**LEVEL 2** *60 points required to complete Level 2: At least 30 points from 2 different quests*	30	20	100
	Quest 1 *50 points* / Quest 2 *50 points* / Quest 3 *50 points*	30	20	100
4 & 5	**LEVEL 3** *60 points required to complete Level 3: At least 30 points from 2 different quests*	30	20	100
	Quest 4 *50 points* / Quest 5 *50 points* / Quest 6 *50 points*	30	20	100
6 & 7	**LEVEL 4** *60 points required to complete Level 4: At least 30 points from 2 different quests*	30	20	100
	Quest 7 *50 points* / Quest 8 *50 points* / Quest 9 *50 points*	30	20	100
8 & 9	**LEVEL 5** *60 points required to complete Level 5: At least 30 points from 2 different quests*	30	20	100
	Quest 10 *50 points* / Quest 11 *50 points* / Quest 12 *50 points*	30	20	100
10 & 11	**LEVEL 6** *60 points required to complete Level 6: At least 30 points from 2 different quests*	30	20	100
	Quest 13 *50 points* / Quest 14 *50 points* / Quest 15 *50 points*	30	20	100
12 & 13	**LEVEL 7** *60 points required to complete Level 7: At least 30 points from Quest 16 and Quest 17*	30	20	100
	Quest 16 *50 points* / Quest 17 *50 points* / Quest 18 *50 points*	30	20	100
	THE END! Maximum points = 650	390	260	1300

The homework

The quests in this study are language learning tasks that the participants completed independent of the teacher. 20 quests were developed, trialled (Philpott, 2015b), and edited prior to the commencement of this study. 18 quests were included in this gamified course. To avoid overwhelming participants with a choice of 18 quests, levels were used to control the delivery of quests. The quests were categorised into levels based on themes and supporting academic theory. Seven levels were created; however, Level 1 was an introduction level that did not have any quests; Level 2 to Level 7 had three quests each. When a participant had completed at least two quests in a level, they could progress to the next level. The participants were told to complete one quest per week. The participants could do all three quests in a level; however, to progress to the end of the levels on schedule it was recommended to only do two quests per level. Quests were assessed each week through a peer review

system. Table 3.3 shows the outline of all the levels, themes, quests, type of submission required, psychological theory, target language skills, number of people required to complete each quest, and points. All the quests were designed to foster the three tenets of SDT; however, other psychological theories were used in conjunction with SDT to target specific aspects of motivation. See Appendix A for a full description of each quest.

Table 3.3

Table of Levels and Quests

Week	Quest Name	Submission Type	Psychological Theory	Target Language Skill	People Required	Points
LEVEL 1: LET'S GET STARTED			Level 1 introduces students to the quests and blogging system.			Max = 50
1	Introduction and blog creation	Write	Onboarding	Internet, Writing	1	50
LEVEL 2: GETTING TO KNOW YOU			Level 2 encourages students to self-reflect and set goals.			Max = 100
2 - 3	Q1: About Myself	NPP	Foster meaningfulness, Goal-setting	Presentation, Pronunciation	1	50
	Q2: Self-Assessment	Write		Writing, Vocabulary	1	50
	Q3: Semester Goals	Video		Speaking	2	50
LEVEL 3: IT'S GOOD FOR YOU			Level 3 encourages meaningful personal and interpersonal experiences.			Max = 100
4 - 5	Q4: A Foreign Friend	Write	Encourage International Posture	Speaking, Listening, Writing	2	50
	Q5: Pay it Forward	Write / Video	Foster relatedness	Writing, Speaking	1 or 2	50
	Q6: 30-day Challenge	Write	Foster challenge	Writing, Vocabulary	1	50
LEVEL 4: LOOKING FORWARD TO THE FUTURE			Level 4 encourages students to think about their future.			Max = 100
6 - 7	Q7: Interesting Jobs	Video	Ideal/Future self, encourage International Posture	Speaking, Listening	2	50
	Q8: Living Abroad	Write		Writing, Vocabulary	1	50
	Q9: My Major	NPP		Presentation, Vocabulary, Pronunciation	1	50
LEVEL 5: THE WORLD AROUND YOU			Level 5 encourages students to think about international issues.			Max = 100
8 - 9	Q10: Video News	Write	Encourage International Posture	Listening, Vocabulary, Writing	1	50
	Q11: Interview	Video		Speaking, Listening, Interview	2	50
	Q12: Ted Talks	Video		Listening, Vocabulary, Pronunciation	1	50
LEVEL 6: LEARNING THROUGH NEW MEDIA			Level 6 encourages learner autonomy.			Max = 100
10 -11	Q13: Documentary Discussion	Video	Encourage autonomy	Speaking, Vocabulary	2-3	50
	Q14: ELLLO	Video / Write		Listening, Speaking, Pronunciation	1, 2, 3+	50
	Q15: Podcasts	Write		Listening, Vocabulary	1	50
LEVEL 7: TIME TO REFLECT			Level 7 encourages reflection and closure to the course.			Max = 100
12 - 13	Q16: Presentation Review	Write	Meaningful reflection	Self-assessment, writing	1	50
	Q17: Reflecting on the Semester	Write		Self-assessment, writing	1	50
	Q18: Design your own Quest	Write	Encourage autonomy	Writing	1	50
				Maximum possible quest points = 650		

To facilitate the extensive reading (explained in Section 1.3) component of the homework, the participants had access to over 1000 graded readers in the host university's library. MReader (http://mreader.org) was used to track each participant's extensive reading progress. When the participants finished reading a book, they logged into the MReader website to do quiz about the book they just read. The quiz asked questions to determine if the participants read and understood the book or not. Each quiz took approximately 10 minutes to complete. Once a participant passed a quiz for a specific book, he/she could not do that quiz again. MReader had built in features that alert the researcher of possible participant collusion. The researcher had access to all the data MReader tracked for each participant. The data most pertinent to this study were the number of quizzes passed in a specific week and the number of quizzes failed in that same specific week. MReader was a free-to-use website.

For the English Central (explained in Section 1.3) component of the homework, the participants' weekly goal was to watch and speak three videos. Speaking a video refers to using English Central's digital speech shadowing program that requires the users to repeat the dialogue they heard in a video. The researcher had access to each participant's performance data. The host university had paid for the participants to have unlimited access to English Central.

The ecosystem

Due to the inability of other gamified learning management systems (introduced in Section 2.3.2) to allow students to easily produce, share, and take ownership of their completed work, the researcher developed a gamified ecosystem that would be suitable for the goals of this study. The digital ecosystem was designed to connect all aspects of the gamified course. Each class had a class website that was created using Google Sites (sites.google.com). The websites detailed all the important information for each class and had links to the quests, the quest scoring form, the blogs, and the leaderboards for Class 1. Blogger (www.blogger.com) was used to deliver quests to the participants. Each quest level had its own blog. The levels were connected using hyperlinks. The participants were told that they could click on the Change Level link once they had successfully completed two quests in a level. Completing a quest required the participant to get at least 30 of the 50 points. Blogger was also used as a

privacy-controlled space for the participants to submit and share their completed quests. Blogger allowed for work to be commented on and allowed for the integration of various types of media such as sound recordings, videos, pictures, and PowerPoint presentations. All of the participants had their own blog, and each week they would do one post on their blog. The post had two sections: (1) a short personal diary entry, and (2) the outcome of the quest they completed. The researcher hoped that the writing and sharing of a personal diary would be another avenue to foster the tenets of SDT through authentic L2/FL use. Figure 3.2 shows how Quest 3 appeared to the participants. Figure 3.3 shows how a participant's blog appeared at the end of the study.

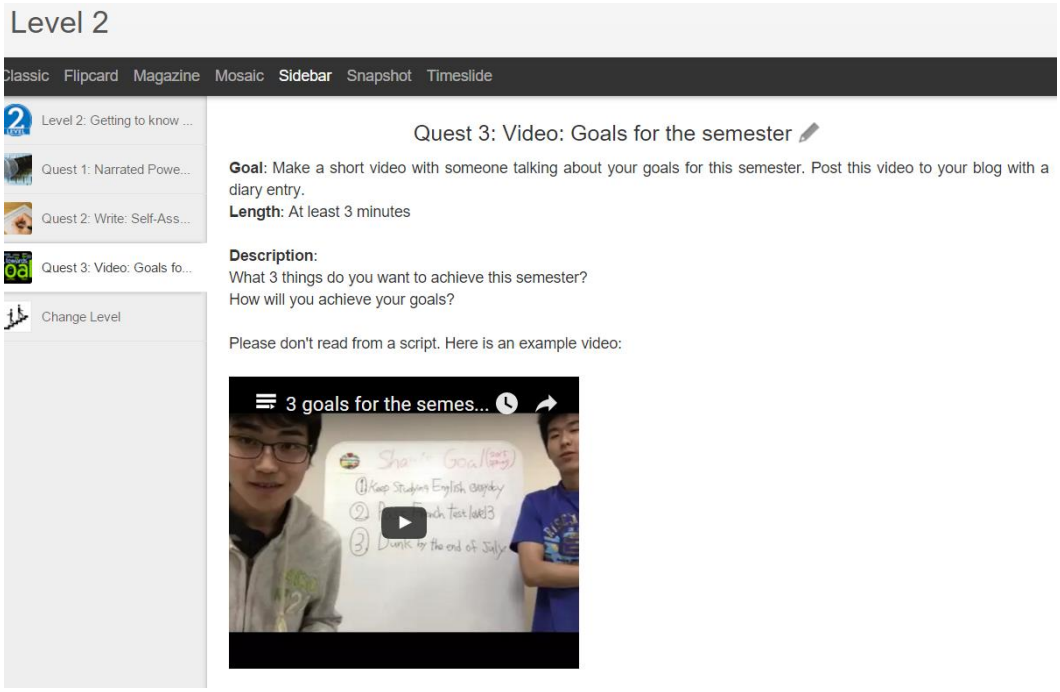


Figure 3.2. Quest 3 screenshot.
Note: This screenshot shows how a quest appeared to the participants.

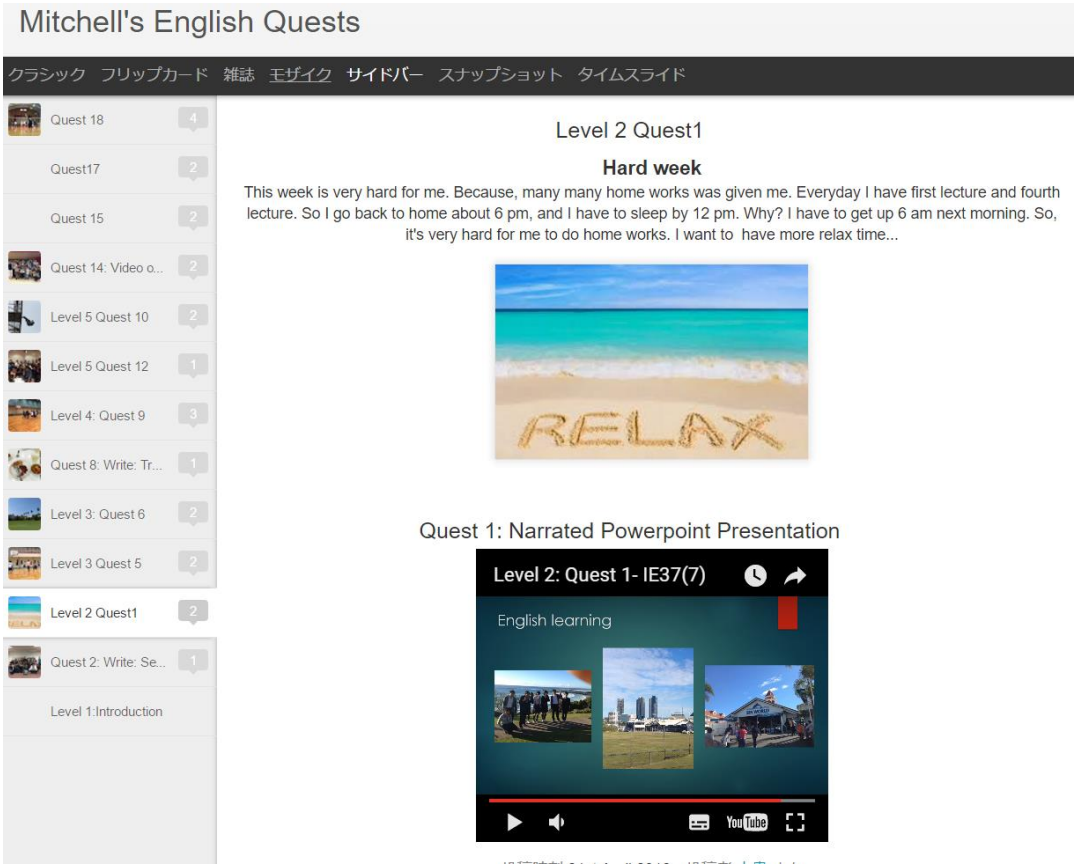
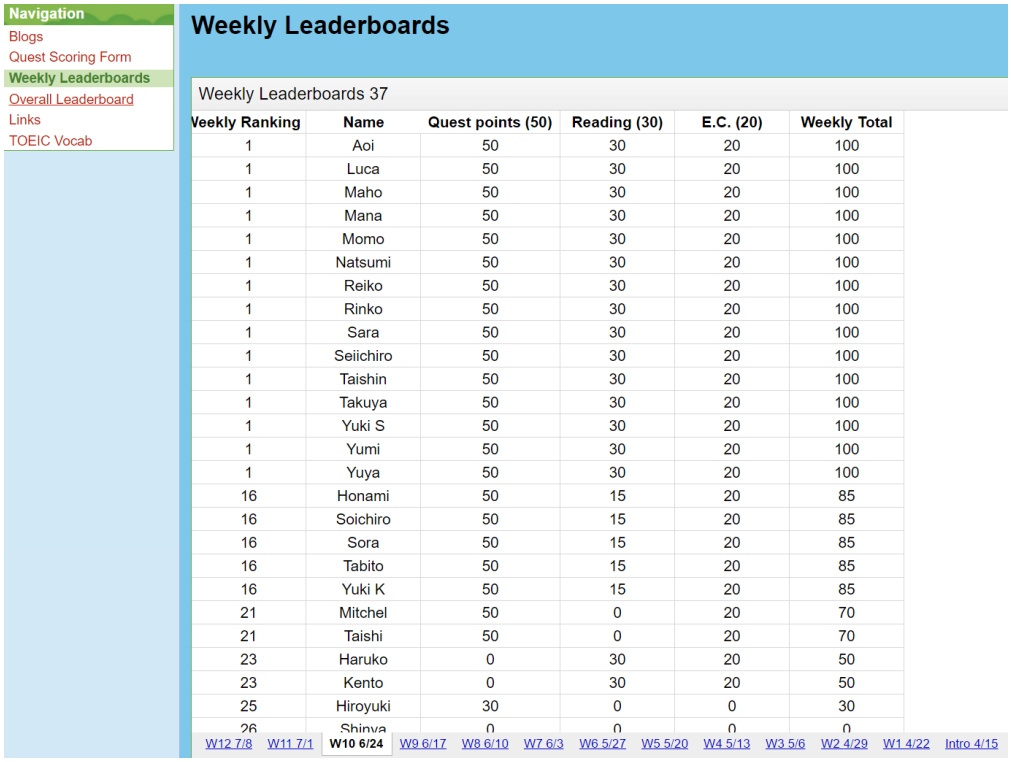


Figure 3.3. A participant's completed quest.
Note: This is a screenshot of a participant's completed quest following their weekly diary entry.

Every Friday in class, a website (http://www.aschool.us/random/random-pair.php) was used to randomly assign each participant two blog posts to check. The participants checked the blog posts by reading the diary entry and then assessing the completed quest based on the quest objective and the point scoring system. The quest scoring system relied on the participants to accurately assess each other's work; the teacher would occasionally check the accuracy of the assessments. Having the participants assess each other served two purposes. First, it ensured that the participants' work would be shared. Second, it greatly decreased the burden on the teacher to assess 46 quests each week. The scores were submitted to the teacher using a Google Form (forms.google.com) that was embedded on the class website. After completing the form, the participants were encouraged to write a comment in the comments section of the blog they assessed. There was no requirement for how long the comment should be or what the comment should be about. However, the participants were encouraged to use this commenting phase as an opportunity for social communication rather than critical feedback. Each quest was graded by two

different participants. The teacher checked for score discrepancies and then input the quest scores into Google Sheets (sites.google.com) spreadsheets alongside each participant's weekly English Central score and extensive reading score. The spreadsheets combined the three scores based on the points scoring system to create a weekly point total and an overall point total for each participant. The spreadsheet sorted the participants from high to low based on their points' totals; this sorted spreadsheet acted as a leaderboard. The leaderboard was embedded on the website for Class 1. The leaderboard was automatically updated once the page was refreshed. Figure 3.4 shows the weekly leaderboard at Week 10. Figure 3.5 shows the overall leaderboard at Week 10.



Navigation
Blogs
Quest Scoring Form
Weekly Leaderboards
Overall Leaderboard
Links
TOEIC Vocab

Weekly Leaderboards

Weekly Leaderboards 37

Weekly Ranking	Name	Quest points (50)	Reading (30)	E.C. (20)	Weekly Total
1	Aoi	50	30	20	100
1	Luca	50	30	20	100
1	Maho	50	30	20	100
1	Mana	50	30	20	100
1	Momo	50	30	20	100
1	Natsumi	50	30	20	100
1	Reiko	50	30	20	100
1	Rinko	50	30	20	100
1	Sara	50	30	20	100
1	Seiichiro	50	30	20	100
1	Taishin	50	30	20	100
1	Takuya	50	30	20	100
1	Yuki S	50	30	20	100
1	Yumi	50	30	20	100
1	Yuya	50	30	20	100
16	Honami	50	15	20	85
16	Soichiro	50	15	20	85
16	Sora	50	15	20	85
16	Tabito	50	15	20	85
16	Yuki K	50	15	20	85
21	Mitchel	50	0	20	70
21	Taishi	50	0	20	70
23	Haruko	0	30	20	50
23	Kento	0	30	20	50
25	Hiroyuki	30	0	0	30
26	Shinya	0	0	0	0

W12 7/8 | W11 7/1 | **W10 6/24** | W9 6/17 | W8 6/10 | W7 6/3 | W6 5/27 | W5 5/20 | W4 5/13 | W3 5/6 | W2 4/29 | W1 4/22 | Intro 4/15

Figure 3.4. Week 10 leaderboard of the participants' weekly points.
Note: This is a screenshot of the class leaderboard which shows each participant's weekly points for Week 10.

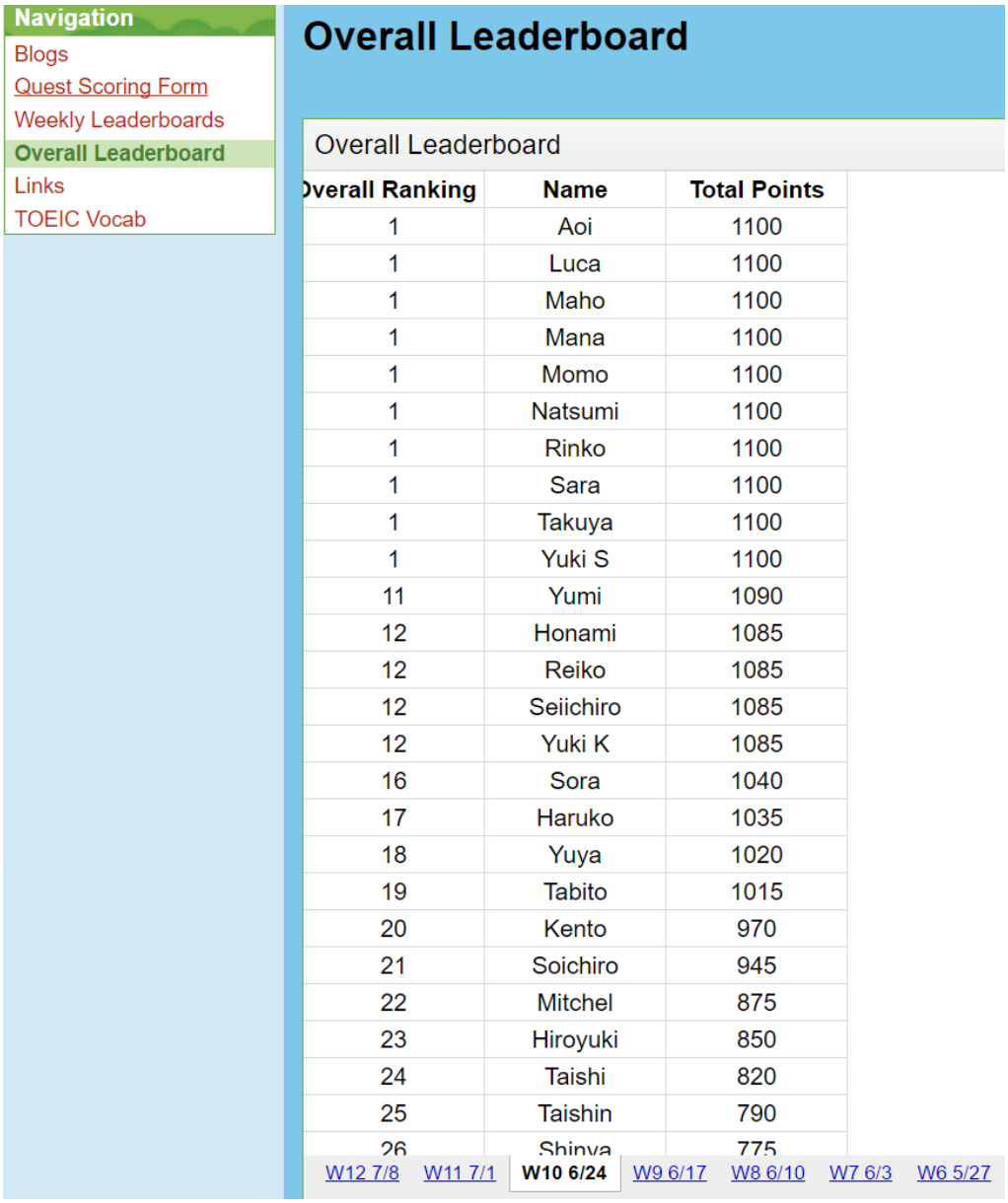


Figure 3.5. Week 10 leaderboard of the participants' overall points.
Note: This is a screenshot of the class leaderboard which shows each participant's overall points for Week 10.

3.5 Data Collection Instruments

Five data collection instruments were used in this study: performance-related data, a leaderboard questionnaire, quest diaries, semi-structured interviews, and the LLOS. The performance-related data and the LLOS collected quantitative data, the leaderboard questionnaire and the quest diaries collected mixed data, and the semi-structured interviews collected qualitative data. The instruments were first created in English and then translated into Japanese by a bilingual graduate student. The

Japanese was then isolated and then back-translated into English by a different bilingual graduate student. Any differences between the original and the back-translated version were resolved through discussion between the two translators to ensure the Japanese translation was accurate. All the instruments were delivered in English with the Japanese written below. The participants could answer in either Japanese or English. However, they were encouraged to use Japanese if they could not clearly express themselves in English.

3.5.1 Performance-Related Data

At the end of each week for 12 weeks, quantitative performance-related data for all the participants were collected from English Central, MReader, and the quest scoring forms, and stored in Excel spreadsheets. The data collected from English Central included the number of videos watched (EC_videos_watched), the number of videos spoken (EC_videos_spoken), and the percentage of English Central goals achieved (EC_goals_completed). The EC_goals_completed variable represented a combination of the EC_videos_watched and EC_videos_spoken variables, and it was used as a component for the leaderboard point system (presented in Table 3.1). The data collected from MReader included the number of quizzes passed (MR_quizzes_passed), the number of words read (MR_words_read). Data collected in relation to the quests were the number of quests completed (Quest_completed). In addition, for Class 1, each participant's weekly leaderboard points and leaderboard rank (LB_rank) was tracked.

A different spreadsheet tracked all the participants' individual accumulated performance data. Table 3.4 presents an excerpt from the final accumulated performance spreadsheet. The number in brackets after each performance measure shows the final target point totals that would be achieved at the end of the 12-week period if the participants successfully completed their weekly goals. The maximum points total was 1300 rather than 1200 as all the participants in Class 1 were awarded 100 points as part of the course introduction. MR_words_read did not have a target goal, and it did not have any bearing on the leaderboard.

Table 3.4
End of Study Performance-Related Data from Two Participants

Participant #	EC_ videos_ watched (36)	EC_ videos _spoken (36)	EC_ goals_ completed (100%)	MR_ quizzes_ passed (12)	MR_ words _read	Quests_ completed (12)	Points	LB_ rank
1	39	38	106.9	15	34389	13	1300	1
18	32	34	91.7	5	53166	12	1135	18

3.5.2 Leaderboard Questionnaire

The leaderboard questionnaire is a self-report questionnaire designed to collect data about three areas of interest: the participants' emotional reactions to seeing their leaderboard ranking (emotion section); the participants' attitudes toward various aspects of the leaderboard (attitude section); and the participants' opinions about the leaderboards (opinion section). The leaderboard questionnaire was developed based on data collected by the researcher (Philpott, 2015a), and the Achievement Emotions Questionnaire (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). The leaderboard questionnaire was completed by the participants in Class 1 at Week 4, Week 8, and Week 12. The researcher chose to administer the questionnaire only every four weeks to avoid possible survey fatigue that could have occurred from more frequent surveying. For the third and final time the questionnaire was administered, some redundant questions were excluded, and some reflective questions were added. The leaderboard questionnaire was administered during class time using Survey Monkey. The questionnaire took approximately five to ten minutes to complete. The Leaderboard questionnaire can be found in Appendix C.

The emotion section of the leaderboard questionnaire asked the participants to report what emotions they felt when they looked at the weekly leaderboard and the overall leaderboard. Data were collected separately about the weekly and overall leaderboard to determine if the participants reacted differently to one or the other based on their ranking. The participants could choose from ten different emotions and they could choose as many emotions as they desired that indicated their emotional reactions to seeing their leaderboard rankings. The emotions that the participants could choose were enjoyment, hope, pride, determined, surprised, anxiety, shame, hopelessness,

envy, and other. The participants were then asked to explain why they chose that emotion.

The attitude section of the leaderboard questionnaire asked the participants to respond to 12 statements about the leaderboards. The participants chose their response from a 6-point Likert scale anchored by 1 (*strongly disagree*) and 6 (*strongly agree*). The 12 statements were:

1. Gain_rank = When I see the leaderboard, I feel I must do more work to gain rank on the leaderboard.
2. Maintain_rank = When I see the leaderboard, I feel I must do work in order to maintain rank on the leaderboard.
3. Don't care = I don't care where my ranking is on the leaderboard.
4. Show_teacher = When I see the leaderboard, I feel I must do work in order to show the teacher that I'm a good student.
5. Show_students = When I see the leaderboard, I feel I must do work in order to show other students that I'm a good student.
6. Prove_to_myself = When I see the leaderboard, I feel I must do work in order to prove to myself that I'm a good student.
7. Avoid_emb = When I see the leaderboard, I feel I must do more work to avoid embarrassment.
8. Improve_grade = When I see the leaderboard, I feel I must do more work to improve my class grade.
9. LB_is_fun = the leaderboard is fun.
10. Activities_neg = The activities (Reading, English Central, Quests) are not enjoyable so I don't care about the leaderboard.
11. Motivated_Eng = When I see the leaderboard, I feel motivated to do more work to improve my English ability.
12. Remind_Eng = The leaderboard reminds me that improving my English ability is important.

Finally, the opinion section of the leaderboard questionnaire was an open-ended question that provided the participants an opportunity to share their opinions about any aspects of the leaderboard or the class if they wanted to.

3.5.3 Quest Diaries

An instrument was required to collect rich data about the participants' opinions and perceptions of QBL. As the concept of questing is relatively new in EFL, the researcher wanted to use a tool which allowed the participants to freely share their opinions. Similarly to other studies (e.g., Bailey & Nunan, 1996; Hatch, 2002; McLeod, 2003; Neale & Flowerdew, 2003), student diaries were chosen for this purpose as they allow the researcher direct insight into the participants' perspectives. The student diaries, referred to as quest diaries, collected data from all the participants about their retrospective thoughts towards the specific quests and QBL in general. The participants wrote in their quest diaries at the end of each level from Level 2 to Level 6; approximately every two weeks. The participants were told that they could write about anything they wanted to write about; however, they were encouraged to write about why they chose to do a quest, and what they thought about the quest.

At the end of Level 7, the participants were prompted to complete a final quest diary. The final quest diary was more of a questionnaire than a diary entry because it asked the participants to respond to 14 open-ended questions that reflected on certain aspects of the course, with a focus on questing. The final quest diary survey was inserted into the bottom of each participant's quest diary during Week 11 of the study. The final quest diary was inserted at this time because the researcher wanted to add, delete, or edit questions based on observed emerging points of interest. To make the final quest diary more aesthetically pleasing and meaningful, a class photo and a personalised message from the researcher was included in the introduction of the final quest diary. The participants completed the quest diaries outside of class time or during the assigned game-time in Friday's class. It took the participants approximately 5 to 20 minutes to complete a quest diary entry.

A quest diary was created for and shared with each participant using Google Docs. The participants could access their quest diary through the shared folder in their Google Drive, or by clicking a hyperlink that was located on the change level page for each quest level (see Figure 3.6). When the participants accessed their quests diaries, they were welcomed with an introduction message that explained all the

relevant information they needed to know to complete their diaries (see Figure 3.7). Each quest diary was formatted in a way that was easy for the participants to understand and follow. The participants could use Japanese or English to write in their quest diaries; however, they were told to use Japanese if they could not express themselves clearly in English.

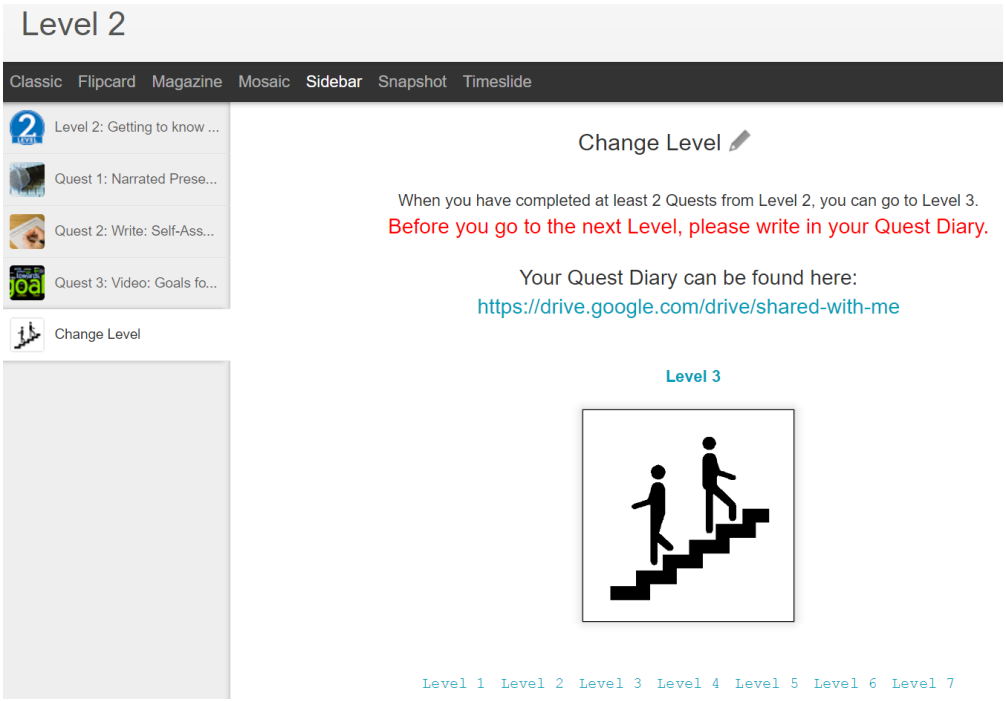
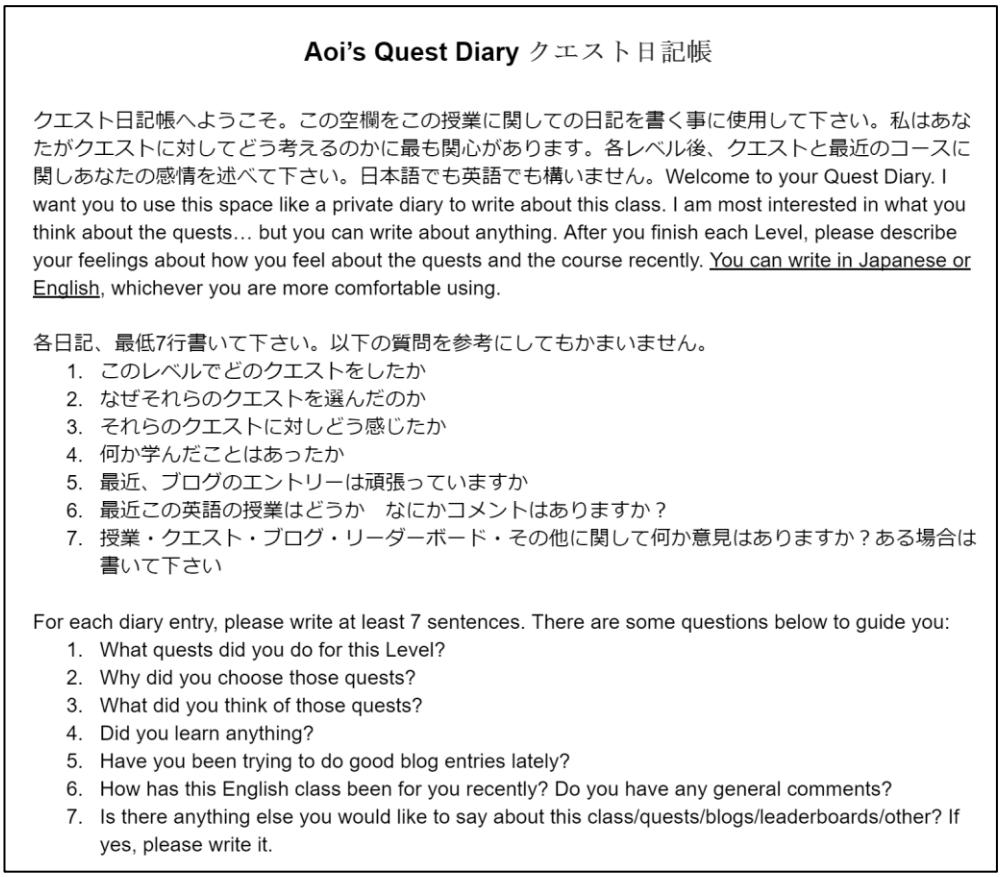


Figure 3.6. Level change and quest diary access.
Note: This is a screenshot of how participants change quest levels and how quest diaries are accessed.



Aoi's Quest Diary クエスト日記帳

クエスト日記帳へようこそ。この空欄をこの授業に関しての日記を書く事に使用して下さい。私はあなたがクエストに対してどう考えるのかに最も関心があります。各レベル後、クエストと最近のコースに関しあなたの感情を述べて下さい。日本語でも英語でも構いません。Welcome to your Quest Diary. I want you to use this space like a private diary to write about this class. I am most interested in what you think about the quests… but you can write about anything. After you finish each Level, please describe your feelings about how you feel about the quests and the course recently. You can write in Japanese or English, whichever you are more comfortable using.

各日記、最低7行書いて下さい。以下の質問を参考にしてもかまいません。

1. このレベルでどのクエストをしたか
2. なぜそれらのクエストを選んだのか
3. それらのクエストに対しどう感じたか
4. 何か学んだことはあったか
5. 最近、ブログのエントリーは頑張っていますか
6. 最近この英語の授業はどうか　なにかコメントはありますか？
7. 授業・クエスト・ブログ・リーダーボード・その他に関して何か意見はありますか？ある場合は書いて下さい

For each diary entry, please write at least 7 sentences. There are some questions below to guide you:

1. What quests did you do for this Level?
2. Why did you choose those quests?
3. What did you think of those quests?
4. Did you learn anything?
5. Have you been trying to do good blog entries lately?
6. How has this English class been for you recently? Do you have any general comments?
7. Is there anything else you would like to say about this class/quests/blogs/leaderboards/other? If yes, please write it.

Figure 3.7. Quest diary introduction page.
Note: This is a screenshot of a quest diary introduction page.

3.5.4 Semi-structured Interviews

Semi-structured interviews were used to expand on and clarify various data already collected throughout the study. The interviews were conducted in Japanese by a Japanese female graduate student studying Applied Linguistics at the host university. The researcher trained the interviewer the protocol for conducting the interviews for this study. Before the interviews were conducted, the researcher prepared an interview guide for the interviewer to use during the interviews. An interview guide was used to provide consistency over the content covered in each interview. As suggested in Corbin and Strauss (2015) and Hatch (2002), the interviewer was also expected to ask clarifying or follow-up questions, or probe into interesting areas that arose during the interview.

Based on a mixture of Patton's (2015) purposeful sampling strategies, participants for the interview were chosen based on availability, willingness to participate, and

various predetermined criteria. The first criterion was choosing participants who represent different types of L2/FL motivation according to the data collected from the LLOS: intrinsic, extrinsic and mixture. The second criterion was choosing participants to represent the final leaderboard ranking categories: high, medium and low. The third criterion was selecting a group of participants that evenly represented males and females. The fourth criterion, based on data collected from the leaderboard questionnaire and the quest diaries, was choosing participants who had different perspectives on the gamified course. All participants in Class 1 were emailed to check their availability and willingness to participate. The participants were told that if they participated, they would receive a 2000-yen Starbucks gift card. 16 participants said they were available to participate; 11 participants said they were available to participate on the most suitable day. Based on the above criteria, semi-structured interviews were conducted with nine participants (n = 9). The interviews lasted for about 15-20 minutes.

3.5.5 The Language Learning Orientation Scale

The *Language Learning Orientations Scale – Intrinsic Motivation, Extrinsic Motivation, and Amotivation Subscales* (LLOS) (Noels, Pelletier, Clément, & Vallerand, 2000) is a psychometric multi-scale questionnaire that measures L2/FL motivation orientation based on SDT. The LLOS comprises 21 statements that measure seven subscales: three subscales entitled *Intrinsic Motivation – Knowledge*, *Intrinsic Motivation – Accomplishment*, and *Intrinsic Motivation – Stimulation* measure different types of intrinsic motivation; three subscales entitled *Identified Regulation*, *Introjected Regulation*, and *External Regulation* measure extrinsic motivation; one subscale entitled *Amotivation* measures amotivation. Each subscale is measured using three statements. The participants completing the LLOS read each statement, decided how much they agreed with the statement, and then assigned their corresponding score using a 7-point scale (*1: does not correspond, 2: corresponds very little, 3: corresponds a little, 4: corresponds moderately, 5: corresponds a lot, 6: corresponds almost exactly, 7: corresponds exactly*). After completing the survey, scores for each subscale were calculated. If a participant gets a high score for a subscale, this suggests that he/she possesses that type of motivation; a low score

suggests the opposite. Based on Kline's (1999) criteria for evaluating internal consistency in which an alpha greater than 0.6 is acceptable, the LLOS (Noels et al., 2000; Vallerand et al., 1992) has demonstrated sufficient reliability; therefore, no major changes were made to it.

The LLOS was administered to all the participants as a pre- and post-test to determine if the participants' FL motivation changed over the duration of the study. The website Survey Monkey (https://www.surveymonkey.com/) delivered the LLOS to the participants. Survey Monkey randomised the order delivery of statements "to create a sense of variety and to prevent respondents from simply repeating previous answers" (Dörnyei, 2010, p. 47). All the participants (N = 46) completed the LLOS outside of class time before Week 1 and after Week 12 of the gamified course. The complete list of questionnaire items is shown in Appendix B.

3.6 Data Collection Procedures

Friday, April 8, 2016, was the first day of class for the second-semester of the Intensive English courses. On this day, the researcher explained the study to the students and distributed a general information sheet (see Appendix D) which explained all pertinent details in Japanese and English. Students could decide whether they wanted to participate or not in the study. Signed informed consent forms were obtained from all students on Wednesday, April 13, 2016. Further details about the informed consent process are discussed in Section 3.8. Data collection began soon after the informed consent was complete. Table 3.5 displays the data collection timetable.

Table 3.5
Data Collection Timetable

Week	**< 1**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**> 12**
LLOS * *both classes*	x													x
Performance-related data * *both classes*		x	x	x	x	x	x	x	x	x	x	x	x	
Leaderboard questionnaire * *only Class 1*					x				x				x	
quest diaries * *both classes*			x		x		x		x		x		x	
Semi-structured interviews (*n* = 9) * *class 1*														x

Data collected from the LLOS and the leaderboard questionnaire were exported from Survey Monkey to the researcher's computer as an SPSS file. Every Friday, the performance-related data collected from English Central, MReader, and the quest scoring forms for each participant was stored in Excel spreadsheets. The semi-structured interviews were recorded using two audio recording devices. The audio recordings were transferred to the researcher's computer after all interviews were complete. The recordings were then sent to the interviewer who transcribed them into Japanese. The interviewer then translated all the qualitative data written in Japanese from the interview scripts, the quest diary entries, and the leaderboard questionnaire to English. The English translations were sent to a different bilingual graduate student who back-translated the texts into Japanese. Any differences between the original Japanese text and the back-translated text were resolved by the two translators through discussion. This process of back-translation ensured the accuracy of the English translations.

3.7 Data Analysis

Table 3.6 outlines the main data analysis procedures that occurred on each data set in relation to the research questions.

Table 3.6
Summary of Data Analysis Procedures

Research question	Data	Analysis
1. What effect do leaderboards have on student performance?	• Performance-related data	• Independent samples *t*-test
	• Leaderboard questionnaire	• Quantitative content analysis on reasons for emotion • Friedman rank sum test to measure attitude change • ANOVA to measure attitude based on leaderboard rank
2. What are students' opinions and perceptions of quest-based learning?	• Quest diaries • Semi-structured interviews	• Quantitative content analysis
3. What are the effects of leaderboards and quest-based learning on L2/FL motivation?	• Language learning orientation scale	• Paired samples *t*-test

Qualitative data analyses were conducted using NVIVO version 12. Thematic quantitative content analysis (Berg & Lune, 2012; Cohen, 2011; Hatch, 2012) was performed on the qualitative data collected from the leaderboard questionnaire, quest diaries, and the semi-structured interviews. The first stage of coding included structural coding, open coding, and sentiment coding. The second stage of coding created axial codes that categorised the open codes based on emerging themes in the data. The large amount of quest diary responses were suitable to be analysed using rank order comparison of frequency (Curtis et al., 2001).

Quantitative data analyses were conducted using SPSS version 25. The specific

dependent variables are outlined in Chapter 4. The performance-related data were analysed using independent samples *t*-tests that measured the performance impact of leaderboards in Class 1 compared to no leaderboards in Class 2. Data from the attitude section of the leaderboard questionnaires were analysed using a Friedman rank sum test that measured attitude change over the duration of the study, and an ANOVA that measured attitude towards the leaderboards based on leaderboard rank. The LLOS data were analysed using paired samples *t*-tests that measured motivation change between the pre- and post-tests.

3.8 Ethical Considerations

All aspects of this thesis were conducted strictly in accordance with the proposal submitted to the University of Southern Queensland's Human Research Ethics Committee and in accordance with the ethical requirements of the host university. Ethical requirements of the host university state that, as participation in any research study is not mandatory, teachers and students should in no way feel pressure to participate. Students who choose not to participate should not be affected in any negative way. Research conducted at the host university must abide by the Personal Information Protection Law for Japan. According to the Personal Information Protection Law in Japan, the collecting entity must describe as fully as possible the purposes of using personal information. The researcher did this by going over the general information sheet, consent forms, and the purpose of the study in detail with all potential participants. Both the researcher's and his supervisor's contact details were provided in case any questions arose after the informed consent form was signed. Additionally, the law states that data must be only used for its intended purposes and not be made available to third parties. Therefore, the researcher has ensured restriction of the participants' data use for this study only, thereby preventing unauthorized third party to use, see, or listen to the data in any way. Furthermore, the law makes clear that the data collector shall take necessary measures to prevent the loss, destruction, or damage of the data. The researcher keeps the hard data in a locked cabinet in his office and the soft data on a password-protected computer. This data will be stored for at least five years after the completion of this study. Finally, Japanese law states that upon request by the

participant, the collector must give access or deliver the personal data. The participants were able to discontinue participation and withdraw their data at any time without consequence. They were able to request the summary of results be sent to them in the informed consent form for the questionnaire.

The possible psychological risk to the participants needed to be addressed for this study. As Class 1 was exposed to a communal leaderboard that showed all participants' ranking within the class, some participants were probably going to feel some degree of anxiety when viewing the leaderboard. Anxiety is common in language classes and at an appropriate level it is considered an affective variable to help facilitate language acquisition (Krashen, 2009). The level of anxiety related to the leaderboard was expected to be lower than other activities in language classes such as presentations or group discussions. As the researcher was studying the participants' emotional reactions to the leaderboard, the research did not want to alter the study to avoid some participants feeling anxious. However, an attempt to minimise the psychological risk at the start of the study was made by allowing the participants to choose a pseudonym to represent themselves on the leaderboard. The majority of the participants chose their first name as their pseudonym. The researcher monitored the participants' stress and anxiety throughout the study through observation of behaviour and by reviewing the collected leaderboard questionnaire and quest diary data. It was noted around the end of the study that one participant was not enjoying their leaderboard ranking. This led to a little social embarrassment for this participant; however, it was not to the extent that intervention was required.

Informed consent was conducted in a culturally senstive manner for Japanese students who were also members of a class that the researcher was teaching. In the first week of class, the researcher explained all aspects of the study using a general information sheet (see Appendix D) that was distributed to all possible participants. After explaining the study, the researcher distributed informed consent forms to all students. The students were encouraged to ask the researcher any questions they had, in person or by email. They were informed that the data collection was not anonymous. They were also told to think about whether they would like to and were willing to participate in this study. In order to avoid coercion, it was clearly emphasised that it was the students' choice as to whether they participate or not and

non-participation would affect them in no negative way. If students did want to participate, they were told to read, sign, and bring their informed consent form to class on the following Wednesday. A folder was left at the front of class on Wednesday and students could put their informed consent form in the file whether they had signed it not. Signed informed consent forms were obtained from all students on Wednesday, April 13, 2016. The blank informed consent form can be found in Appendix E. Signed informed consent forms were stored in a locked cabinet in the researcher's office. The general information sheet and the informed consent forms were translated into Japanese to make sure the students clearly understood the study and what they were expected to do.

3. 9 Summary

To examine the use of leaderboards and quests in an EFL course at a Japanese university, this study used a quasi-experimental mixed methods research design. A gamified course was designed to motivate the participants to complete their homework and increase their intrinsic motivation towards studying English. Leaderboards and quests were the main gamification components used to achieve these goals; however, points and levels provided support. Two classes participated in this study. Both classes completed quests; however, the leaderboard was only present in Class 1. The five different data collection instruments used in this study collected data about L2/FL motivation orientation, performance, emotions and attitudes towards leaderboards, and opinions and perceptions towards quests. This chapter has also presented a summary of the data analysis procedures conducted on each data set. This chapter concluded by detailing the legal and ethical requirements of the universities involved in this study.

Chapter 4. Results

4.1 Overview

This chapter presents the results of the data analyses for each data collection instrument. Section 4.2 presents the results of the independent samples *t*-tests that compare the performance-related data for Class 1 and Class 2 for significant differences. Section 4.3 presents the results of the various data analysis procedures that were conducted on the data collected from the leaderboard questionnaire. Section 4.4 details the results of the quantitative thematic content analyses that were performed on the quest diaries. Section 4.5 presents the findings from the semi-structured interviews. Section 4.6 presents the results of the paired samples *t*-tests that analyse the participants' pre- and post-test mean scores for each LLOS subscale to determine if any significant changes occurred. For each instrument, the specific data analysis procedures are detailed before the results are presented. Section 4.7 provides a summary of all the results presented in the chapter.

4.2 Performance-Related Data Independent Samples *t*-Test Results

Independent samples *t*-tests were conducted to compare the mean scores of the performance-related data of Class 1 and Class 2. Six variables were measured: EC_videos_watched, EC_videos_spoken, EC_goals_completed, MR_quizzes_passed, MR_words_read, and Quests_completed. Table 4.1 displays the mean scores for each variable for each class.

Table 4.1
Comparison of Performance-Related Data Mean Scores

	Class	n	*M*	*SD*	SEM
EC_videos_watched (36)	1	26	37.88	8.67	1.70
	2	20	39.35	13.51	3.02
EC_videos_spoken (36)	1	26	35.30	7.65	1.50
	2	20	38.10	13.03	2.91
EC_goals_completed (100%)	1	26	101.98	20.52	4.02
	2	20	107.15	36.64	8.19
MR_quizzes_passed (12)	1	26	9.92	3.68	.72
	2	20	11.50	4.24	.95
MR_words_read	1	26	44349.38	24359.72	4777.33
	2	20	53726.25	17187.39	3843.21
Quests_completed (12)	1	26	11.69	.97	.19
	2	20	11.95	.22	.05

Note. EC = English Central, MR = MReader. The numbers in brackets refer to the maximum score that the participants could get leaderboard points for at the end of the study.

Table 4.2 shows the results of the independent samples *t*-tests. For Table 4.2, The Levene test guides which row to follow. If the probability value in the Sig. column is statistically significant ($p > 0.05$), then variances are assumed unequal and the researcher should use the second row of data (Cohen, Mannion, & Morrison, 2011). The results of the independent samples *t*-test showed no significant differences in the mean scores between Class 1 and Class 2. This suggests that the presence of the leaderboard in Class 1 did not significantly affect performance.

Table 4.2

Independent Samples t-Test Results Comparing Class Performance

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
EC_vids watched	Equal variances assumed	.97	.32	-.44	44	.65	-1.46	3.27	-8.07	5.14
	Equal variance not assumed			-.42	30.61	.67	-1.46	3.46	-8.54	5.60
EC_vids spoke	Equal variances assumed	2.62	.11	-.90	44	.36	-2.79	3.07	-8.98	3.39
	Equal variance not assumed			-.85	28.87	.40	-2.79	3.27	-9.49	3.91
EC_goals completed	Equal variances assumed	2.86	.09	-.60	44	.54	-5.16	8.51	-22.32	11.99
	Equal variance not assumed			-.56	28.04	.57	-5.16	9.12	-23.86	13.53
MR_quizzes passed	Equal variances assumed	.00	.981	-1.34	44	.18	-1.57	1.17	-3.93	.78
	Equal variance not assumed			-1.32	37.75	.19	-1.57	1.19	-3.99	.84
MR_words_ read	Equal variances assumed	4.32	.04	-1.46	44	.15	-9377	6412	-22299	3545
	Equal variance not assumed			-1.52	43.73	.13	-9377	6131	-21736	2982.2
Quests completed	Equal variances assumed	16.16	.00	-1.16	44	.25	-.25	.22	-0.7	.18
	Equal variance not assumed			-1.31	28.39	.20	-.25	.19	-0.66	.14

The difference between the mean score for Class 1 on the variable EC_vids_watched (*M* = 37.88, *SD* = 8.67) and that of Class 2 (*M* = 39.35, *SD* = 13.51) was not statistically significant (*t* = -.447, *df* = 44, two-tailed *p* = .657). The mean score for Class 1 on the variable EC_vids_spoken (*M* = 35.30, *SD* = 7.65) did not differ statistically significantly (t = -.909, df = 44, two-tailed *p* = .368) from that of Class 2 (*M* = 38.10, *SD* = 13.02). The mean score for Class 1 on the variable EC_goals_completed (*M* = 101.98, *SD* = 20.52) did not differ statistically significantly (t = -.607, df = 44, two-tailed *p* = .547) from that of Class 2 (*M* =

107.15, *SD* = 36.64).

The mean score for Class 1 on the variable MR_quizzes_passed (*M* = 9.92, *SD* = 3.68) did not differ statistically significantly (t = -.1.346, df = 44, two-tailed *p* = .185) from that of Class 2 (*M* = 11.50, *SD* = 4.24). The mean score for Class 1 on the variable MR_words_read (*M* = 44349.38, *SD* = 24359.72) did not differ statistically significantly (t = -1.52, df = 43.73, two-tailed *p* = .133) from that of Class 2 (*M* = 53726.25, *SD* = 17187.38). The mean score for Class 1 on the variable Quests_completed (*M* = 11.69, *SD* = .97) did not differ statistically significantly (t = -1.31, df = 28.39, two-tailed *p* = .201) from that of Class 2 (*M* = 11.95, *SD* = .22).

To further explore the performance-related data, Table 4.3 compares the data of both classes based on performance. The participants classified as high performers were the participants who recorded performance results over 20% of what was required for a perfect score; two participants (8%) in Class 1, and six participants (30%) in Class 2 were classified as high performers. The participants classified as middle performers were the participants who recorded performance results between 80-120% of what was required for a perfect score; 18 participants (69%) in Class 1, and 10 participants (50%) in Class 2 were classified as middle performers. The participants classified as low performers were the participants from Class 1 and Class 2 who had not achieved 80% of what was required for a perfect score; six participants (23%) in Class 1, and four participants (20%) in Class 2 were classified as low performers. Table 4.3 shows that the low performing group for both classes had a similar impact on dragging down the class scores. Class 2 had a larger % of middle performers whereas Class 1 had a larger % of high performers. Table 4.4 presents the % of points received during three-week periods for Class 1 based on final leaderboard rank.

Table 4.3
Class Comparison of Performance Data by Group

	Class 1			Class 2		
	group	*M*	*SD*	group	*M*	*SD*
EC_Videos_ watched (36)	high (*8%*)	53.50	4.95	high (*30%*)	53.00	15.32
	mid (*69%*)	40.11	4.48	mid (*50%*)	37.00	3.89
	low (*23%*)	26.00	4.24	low (*20%*)	24.75	6.65
EC_videos_ spoken (36)	high	52.00	7.07	high	52.16	13.46
	mid	36.78	3.14	mid	35.20	4.13
	low	25.33	3.26	low	24.25	6.5
EC_goals_ completed (100%)	high	146.50	16.69	high	146.07	39.25
	mid	105.57	9.23	mid	99.44	10.62
	low	76.40	11.72	low	68.05	18.18
MR_quizzes _passed (12)	high	13.50	3.54	high	14.50	4.50
	mid	11.00	2.54	mid	11.60	2.59
	low	5.5	3.27	low	6.75	3.5
MR_words read	high	65352	11270	high	58823	15310
	mid	49786	22777	mid	54913	13981
	low	21038	15577	low	43112	26209
Quests_ completed (12)	high	12.00	0	high	12.00	0
	mid	12.00	.77	mid	12.00	0
	low	10.66	1.03	low	11.75	.50

Table 4.4
% of Possible Points Received for Class 1

	week 1 - 3	week 4 - 6	week 7 - 9	week 10 - 12
high (*n* = 9)	100	100	100	100
mid (*n* = 9)	83	82	82	76
low (*n* = 8)	92	69	78	67

4.3 Leaderboard Questionnaire Results

The participants in Class 2 (*n* = 26) were shown the weekly and overall leaderboards each week for 12 weeks. The leaderboard questionnaire was administered at Week 4 (Survey 1), Week 8 (Survey 2), and Week 12 (Survey 3) using Survey Monkey. The

leaderboard questionnaire comprised three sections: emotions, attitudes, and opinions. The emotions section collected quantitative and qualitative data about the self-reported emotions the participants felt when they saw the leaderboards. The attitudes section collected quantitative data about the participants' attitudes towards seeing the class leaderboards. The opinions section collected qualitative data about the participants' opinions towards using a leaderboard in class. The quantitative data were analysed using SPSS 25 and the qualitative data were analysed using NVIVO 12.

4.3.1 Leaderboard Questionnaire: Emotions Section Descriptive Statistics

The emotions section asked the participants to look at the weekly and overall leaderboards before choosing an emotion that represented how they felt about each of the leaderboards. The participants could choose from ten emotions, and they could choose as many emotions as they felt necessary to express their range of feelings. Table 4.5 presents the number of times that each emotion was chosen for each survey. Table 4.5 also presents the mean scores for each emotion for each survey and shows the frequency rank compared to the other emotions. The participants who chose the emotion category *other* generally wrote a comment such as 'nothing special'.

Table 4.5
Frequency Table of Emotions

emotion		S1 *n* = 26	S2 *n* = 25	S3 *n* = 26	emotion *M*	rank
enjoyment	weekly	9	7	9	8.33	2
	overall	10	7	9	8.67	
	survey *M*	9.5	7	9	8.50*	
hope	weekly	1	5	2	2.67	4
	overall	2	4	2	2.67	
	survey *M*	1.5	4.5	2	2.67*	
pride	weekly	1	3	3	2.33	5
	overall	2	3	3	2.67	
	survey *M*	1.5	3	3	2.50*	
determined	weekly	22	16	11	16.33	1
	overall	22	16	10	16.00	
	survey *M*	22	16	10.50	16.17*	
surprised	weekly	0	0	2	.67	9
	overall	0	0	2	.67	
	survey M	0	0	2	.67*	
anxiety	weekly	2	6	4	4.00	3
	overall	1	6	2	3.00	
	survey *M*	1.5	6	3	3.50*	
shame	weekly	3	3	1	2.33	7
	overall	2	2	0	1.33	
	survey *M*	2.5	2.5	0.50	1.83*	
hopelessness	weekly	0	2	0	.67	8
	overall	0	2	1	1.00	
	survey *M*	0	2	0.50	0.83*	
envy	weekly	0	0	0	.00	10
	overall	0	0	0	.00	
	survey *M*	0	0	0	.00*	
other	weekly	1	3	2	2.00	5
	overall	1	3	5	3.00	
	survey *M*	1	3	3.50	2.50*	
combined survey *m* total		39.50	44.00	34.00	39.17	

Note. S = survey, for example S1 = Survey 1. * = rank based on this number

The results in Table 4.5 show that the most frequently selected emotion was determined. The mean score for determined (*M* = 16.17) shows that on average, about 16 participants felt determined when they looked at one of the leaderboards.

The second most frequently selected emotion was enjoyment. The mean score for enjoyment (*M* = 8.50) shows that on average, about nine participants felt enjoyment when they looked at one of the leaderboards. The mean scores for the other emotions are listed from high to low: anxiety (*M* = 3.50), hope (*M* = 2.67), pride (*M* = 2.50), other (*M* = 2.50), shame (*M* = 1.83), hopelessness (*M* = .83), surprised (*M* = .67), envy (*M* = .00). Figure 4.1 provides a visual representation of the emotion frequency data.

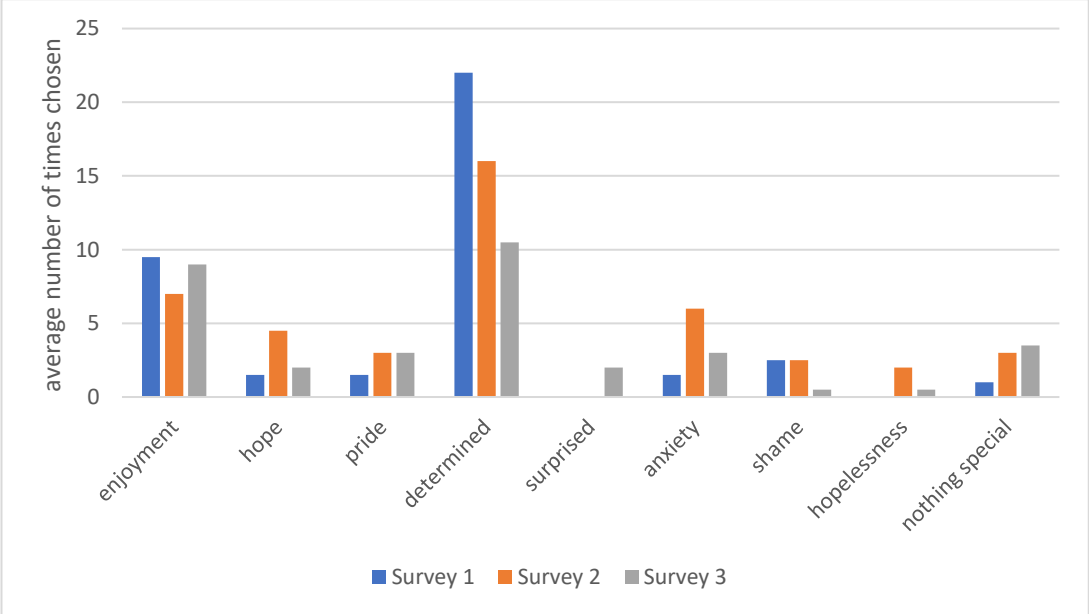


Figure 4.1. Mean frequency for each emotion for each survey.

4.3.2 Reasons Why an Emotion Was Chosen: Content Analysis Results

After the participants chose their corresponding emotions, they explained their choice. A quantitative thematic content analysis was performed on the responses using NVIVO 12. In total, 165 codes were attached to the responses from Survey 1 (*n* = 26), Survey 2 (*n* = 25), and Survey 3 (*n* = 26). Four major thematic categories of responses emerged from the content analysis: rank, personal feeling, score, and social comparison. Each category comprises multiple sub nodes that represent similar themed and more specific responses. Table 4.6 shows the frequency of the most common nodes.

Table 4.6
Reasons Why Any Emotion Was Chosen Content Analysis Results Table

thematic category	node	*freq* = 165
rank (*freq* = 56)	Motivated by rank	42
	Negative feeling about rank	14
personal feeling (*freq* = 47)	Felt sense of accomplishment	26
	Motivated to do work	10
	Negative personal feeling	9
score (*freq* = 32)	Positive opinion about score	27
	Negative opinion about score	5
social comparison (*freq* = 19)	Negative social comparison	14
	Positive social comparison	5

Note. *freq* = frequency.

The *rank* category represents the positive and negative comments the participants had towards the ranking system. Positive comments were generally that the leaderboard motivated them to gain or maintain rank. Negative comments were generally that a low ranking was demotivating because it was difficult to gain rank. The *personal feeling* category represents two positive nodes and one negative node. The positive nodes were that the participants felt a sense of accomplishment or felt motivated to work when they looked at the leaderboard. The negative personal feeling node represents comments about the participants not being able to finish their work or they lacked confidence to perform well.

The *score* category represents the positive and negative comments towards the scoring system. Positive comments were that the scoring system was motivating, and it was useful to track progress. The negative comments were that the scoring system made the participants feel worried or demotivated. The *social comparison* category represents the positive and negative comments about the social comparison aspect of the leaderboards. Negative comments were about the participants feeling worried or ashamed that their score or rank was visible to all participants. Positive comments were about the participants feeling motivated to work because they could see that the other participants were working hard.

The frequency relationships between the nodes and the corresponding emotions were then analysed using a matrix coding query in NVivo 12. As each emotion could have been connected to multiple nodes, the frequency of coded nodes rises from 165 presented in Table 4.6 to 235 in Table 4.7. Table 4.7 displays the results of the analysis that show the most frequent responses as to why a participant chose an emotion and its corresponding percentage weight and ranking. The percentage was calculated to clearly show the importance of the node within the emotion and overall compared to the other emotions.

Table 4.7 shows that the most frequent reasons for choosing the determined emotion were because they were motivated to gain or maintain rank, they felt positively about the scoring system, and they felt a sense of accomplishment. The most frequent reasons for choosing the enjoyment emotion were because they were motivated to gain or maintain rank, they felt a sense of accomplishment, and they felt positively about the scoring system. The most frequent reasons for choosing the anxiety emotion were because they had negative feelings about their rank, and they had negative feelings about other participants seeing their rank or low score. The most frequent reasons for choosing hope were because they felt positively about the scoring system, and they felt a sense of accomplishment. The most frequent reasons for choosing pride were because they felt a sense of accomplishment, they were motivated by rank, and felt positively about the scoring system. The most frequent reason for choosing shame was because they had negative feelings about other participants seeing their rank or low score.

Table 4.7
Frequency Relationship between Emotion and Reason

	freq	within emotion	Overall	
	(235)	%	%	rank
Determined	98		41.70	
Motivated by rank	34	34.69	14.47	1
Positive opinion about score	22	22.45	9.36	2
Felt sense of accomplishment	17	17.35	7.23	3
Motivated to do work	7	7.14	2.98	8
Negative social comparison	7	7.14	2.98	8
Enjoyment	51		30.90	
Motivated by rank	17	33.33	7.23	3
Felt sense of accomplishment	13	25.49	5.53	5
Positive opinion about score	8	15.69	3.40	7
Motivated to do work	6	11.76	2.55	11
Positive: others are working hard, I	4	7.84	1.70	17
Anxiety	22		9.36	
Negative feeling about rank	6	27.27	2.55	11
Negative social comparison	5	22.73	2.13	14
Negative personal feeling	3	13.64	1.28	19
Negative opinion about score	3	13.64	1.28	19
Positive opinion about score	3	13.64	1.28	19
Hope	16		6.81	
Positive opinion about score	5	31.25	2.13	14
Felt sense of accomplishment	4	25.00	1.70	17
Motivated by rank	2	12.50	.85	22
Motivated to do work	2	12.50	.85	22
Pride	20		8.51	
Felt sense of accomplishment	9	60.00	3.83	6
Motivated by rank	6	30.00	2.55	11
Positive opinion about score	5	33.33	2.13	14
Shame	11		6.66	
Negative social comparison	7	58.33	2.98	8
Negative feeling about rank	2	16.67	.85	22

4.3.3 Emotion Activation: Cross-Tabulation Results

To determine if the selection of a specific emotion led to activated behaviour, the emotion data collected from the Leaderboard questionnaires was cross-tabulated with the performance-related data. For Survey 1 and Survey 2, if a participant reported a specific emotion in response to the weekly or overall leaderboard, then that emotion was marked as present for that participant for that survey. A maximum of six different emotions could be registered for each participant for each survey; however, most participants registered only two unique emotions. Activated behaviour was determined based on the performance-related data for the weeks that immediately followed Survey 1 and Survey 2 (i.e., Week 5 and Week 9). If a participant received a maximum score of 100 points, it was considered activated behaviour and they were marked as *max score*, if a participant's score was under 100 points, it was considered not activated behaviour and they were marked as *not max score*. The number of times that an emotion was marked present for the participants that received a max score in Week 5 and Week 9 was calculated.

Table 4.8 presents the activation data for each emotion for Survey 1 and Survey 2 and shows the mean activation percentages for Survey 1 and Survey 2 combined. The emotion data for Survey 3 was omitted as no performance-related data was collected in the week following Survey 3. Table 4.8 shows that for Week 5, 21 of the 26 participants registered a maximum score; for Week 9, 19 of the 26 participants registered a maximum score.

Table 4.8
Emotions and Following Week Performance Table

emotion	S1 *n* = 26	Week 5 max score *n* = 21	S2 *n* = 25	Week 9 max score *n* = 19	S1&S2 mean max score %
enjoyment	10	10	8	7	94%
hope	2	2	6	5	92%
pride	2	2	3	3	100%
determined	22	18	17	15	85%
surprised	0	0	0	0	*NA*
anxiety	2	2	7	4	79%
shame	3	0	3	0	0%
hopelessness	0	0	3	1	33%
envy	0	0	0	0	*NA*
nothing special	1	1	4	2	75%

Note. S = survey, for example S1 = Survey 1.

4.3.4 Emotion Frequency Based on Leaderboard Rank Results

To further explore the emotion data from Section 4.3.1, the emotion frequencies were analysed based on the participants' final leaderboard rank and score. The participants were classified as being either high, medium, or low ranked students. Nine participants with a perfect final leaderboard score of 1300 were ranked as high, nine participants with a score between 1200 to 1299 were ranked as middle, and eight participants with a score less than 1200 were ranked as low. One participant in the middle group did not complete Survey 2, and the low ranked group had a lower number of participants; therefore, the mean percentages were calculated for each group. Table 4.9 displays the mean percentages for each emotion, for each survey, based on the participants' final leaderboard ranking. Figure 4.2 provides a visual representation of the emotion frequency by leaderboard rank data.

Table 4.9

Emotion Frequency Based on Final Leaderboard Ranking Table

LB rank		Enjoyment				Hope			
	n	S1%	S2%	S3%	*M%*	S1%	S2%	S3%	*M%*
high	9	55	55	55	55	11	44	22	26
middle	9	33	25	55	38	0	0	0	0
low	8	25	25	12	21	12	25	0	12
		Pride				Determined			
	n	S1%	S2%	S3%	*M%*	S1%	S2%	S3%	*M%*
high	9	11	22	33	22	88	88	55	77
middle	9	11	12	0	8	88	87	66	80
low	8	0	0	0	0	75	12	25	37
		Surprised				Anxiety			
	n	S1%	S2%	S3%	*M%*	S1%	S2%	S3%	*M%*
high	9	0	0	0	0	11	11	11	11
middle	9	0	0	22	7	0	12	11	8
low	8	0	0	12	4	12	62	25	33
		Shame				Hopelessness			
	n	S1%	S2%	S3%	*M%*	S1%	S2%	S3%	*M%*
high	9	0	0	0	0	0	0	0	0
middle	9	0	12	11	7	0	0	0	0
low	8	37	25	12	25	0	37	12	16
		Envy				Nothing special			
	n	S1%	S2%	S3%	*M%*	S1%	S2%	S3%	*M%*
high	9	0	0	0	0	0	0	0	0
middle	9	0	0	0	0	0	25	11	12
low	8	0	0	0	0	12	25	37	24

Note. S = survey, for example S1 = Survey 1. LB = leaderboard.

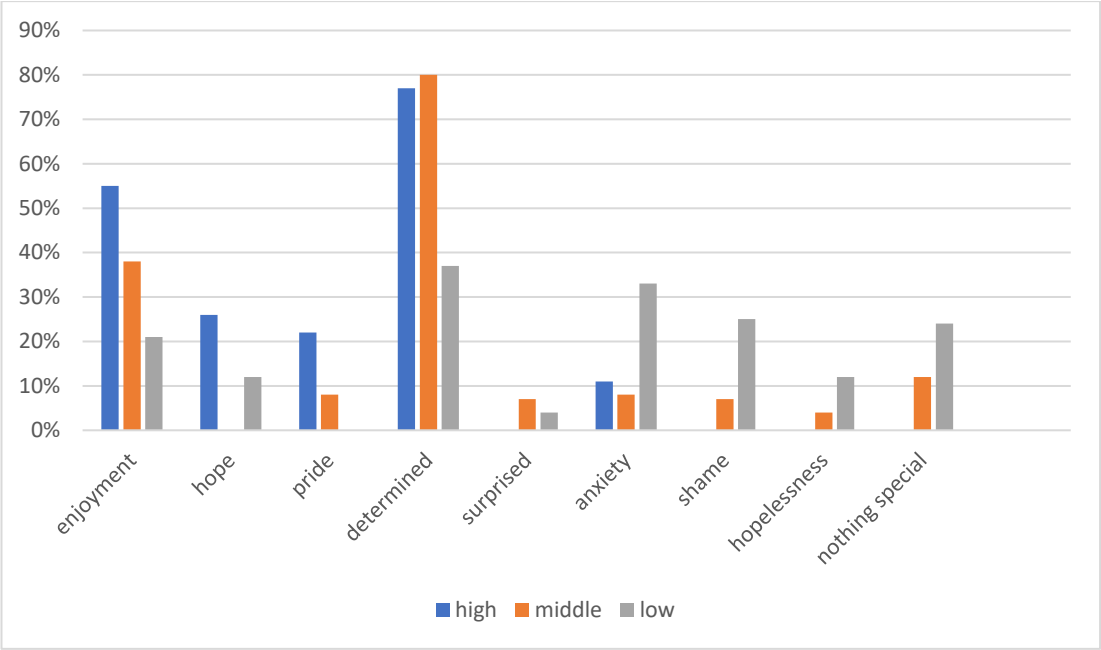


Figure 4.2. Emotion chosen based on final leaderboard rank.

4.3.5 Leaderboard Questionnaire: Attitude Section Descriptive Statistics

The attitude section of the leaderboard questionnaire asked the participants to respond to 12 statements about the class leaderboards. The participants choose their response from a 6-point Likert scale anchored by 1 (strongly disagree) and 6 (strongly agree). Table 4.10 displays the mean scores for each statement, for each survey, and displays each statement's overall ranking compared to the other statements. See Section 3.5.2 for more detailed information about each statement.

Table 4.10
Mean Scores and Overall Ranking for Each Attitude Statement

Attitude Statement	Survey 1 (n = 26)			Survey 2 (n = 25)			Survey 3 (n = 26)			Overall		
	M	*SD*	r	*M*	*SD*	r	*M*	*SD*	r	*M*	*SD*	r
Gain_rank	4.76	1.26	3	4.36	1.28	2	4.40	1.15	3	4.50	0.22	3
Maintain_rank	4.84	1.40	2	4.64	1.28	1	4.56	1.26	1	4.68	0.14	1
Don't_care *	2.68	1.46	11	2.76	1.20	11	3.36	1.75	10	2.93	0.37	11
Show_teacher	4.32	1.41	8	3.60	1.23	9	3.92	1.22	8	3.95	0.36	9
Show_student	3.64	1.78	10	3.08	1.19	10	3.16	1.46	11	3.29	0.30	10
Prove_to_myself	4.44	1.42	5	3.80	1.35	8	3.84	1.25	9	4.03	0.36	8
Avoid_emb	4.04	1.57	9	4.08	1.32	5	4.00	1.38	6	4.04	0.04	7
Improve_grade	4.88	1.27	1	4.36	1.22	2	4.44	1.23	2	4.56	0.28	2
LB_is_fun	4.44	1.08	5	4.04	1.27	6	4.00	1.32	6	4.16	0.24	6
Activities_neg *	2.52	1.23	12	2.48	1.26	12	2.76	1.42	12	2.59	0.15	12
Motivated_Eng	4.56	1.12	4	4.16	1.37	4	4.12	1.42	5	4.28	0.24	4
Remind_Eng	4.36	1.32	7	3.96	1.37	7	4.16	1.31	4	4.16	0.20	5
Combined + *M*	4.42			4.00			4.06			4.16		
Combined – *M* *	2.60			2.62			3.06			2.76		

Note. r = rank, * signifies that the statement represents a negative construct

The results presented in Table 4.10 show that the participants generally agree that, when they saw the leaderboard, they felt they should do more work to maintain their rank (M = 4.68, r = 1); they should do more work to improve their class grade (M = 4.56, r = 2); they should do more work to gain rank (M = 4.50, r = 3); they felt motivated to do more work to improve their English ability (M = 4.28, r = 4); they were reminded that improving their English ability is important (M = 4.16, r = 5); they felt that the leaderboard was fun (M = 4.16, r = 6); they should do more work to avoid embarrassment (M = 4.04, r = 7); they should do more work to prove to themselves they are good students (M = 4.03, r = 8); and they should do more work to show the teacher that they are good students (M = 3.95, r = 9).

The results also show that the participants slightly agree or slightly disagree that, when they see the leaderboard, they should do more work to show the other participants that they are good students (M = 3.29, r = 10), they don't care where their ranking is (M = 2.97, r = 11), and they don't enjoy the learning activities so they

don't care about the leaderboard (M = 2.59, r = 12). The Combined + statement represents the amalgamation of the 10 positively nuanced attitudes, and the Combined – statement represents the amalgamation of the two negatively nuanced attitudes. The results show that on average, the positive attitudes declined between Survey 1 and Survey 2, whereas the negative attitudes increased between Survey 2 and Survey 3.

4.3.6 Measuring Attitude Change: Friedman Rank Sum Test Results

To determine how the participants' attitudes towards each statement changed over the three surveys, Friedman rank sum tests were conducted on each of the 12 statements to examine the equality of their median scores for S1, S2, and S3. The Friedman test is a non-parametric alternative to the repeated measures one-way ANOVA and does not share the ANOVA's distributional assumptions (Zimmerman & Zumbo, 1993). Table 4.11 shows the results of the tests. The results were significant for three statements: Don't_care, Prove_to_myself, and Improve_grade. There were no significant results for the remaining tests, indicating their values were similar: Gain_rank ($\chi^2(2) = 2.926$, $p = .232$), Maintain_rank ($\chi^2(2) = 1.754$, $p = .416$), Show_teacher ($\chi^2(2) = 4.827$, $p = .090$), Show_students ($\chi^2(2) = 3.155$, $p = .206$), Avoid_emb ($\chi^2(2) = .187$, $p = .911$), LB_is_fun ($\chi^2(2) = 2.844$, $p = .241$), Activities_neg ($\chi^2(2) = 1.069$, $p = .586$), Motivated_Eng ($\chi^2(2) = 2.471$, $p = .291$), and Remind_ Eng ($\chi^2(2) = 2.22$, $p = .195$).

For Don't_care, the result $\chi^2(2) = 8.935$, $p = .011$ indicates significant differences in the median values of Don't_care S1, Don't_care S2, and Don't_care S3. For Prove_to_myself, the result $\chi^2(2) = 7.658$, $p = .022$ indicates significant differences in the median values of Prove_to_myself S1, Prove_to_myself S2, and Prove_to_myself S3. For Improve_grade, the result $\chi^2(2) = 6.035$, $p = .049$ indicates significant differences in the median values of Improve_grade S1, Improve_grade S2, and Improve_grade S3. Dunn-Bonferroni post hoc tests (Marshall & Marquier, n.d.) were carried out on the significant results to determine where the difference occurs. Statistically significant results were found between Don't_care S1 and Don't_care S3 ($p = 0.011$), and Don't_care S2 and Don't_care S3 ($p = .004$) after

Bonferroni adjustments. There were no significant differences between any other variables.

Table 4.11
Friedman Rank Sum Test Results That Measured Attitude Change towards the Leaderboards

Variable	Mean Rank	χ^2	*df*	*p*
Gain_rank S1	2.20	2.926	2	>.05
Gain_rank S2	1.86			
Gain_rank S3	1.94			
Maintain_rank S1	2.16	1.754	2	>.05
Maintain_rank S2	1.98			
Maintain_rank S3	1.86			
Don't_care S1	1.86	8.985	2	<.05**
Don't_care S2	1.76			
Don't_care S3	2.38			
Show_teacher S1	2.30	4.827	2	>.05
Show_teacher S2	1.78			
Show_teacher S3	1.92			
Show_students S1	2.24	3.155	2	>.05
Show_students S2	1.84			
Show_students S3	1.92			
Prove_to_myself S1	2.38	7.658	2	<.05
Prove_to_myself S2	1.72			
Prove_to_myself S3	1.90			
Avoid_emb S1	2.06	0.187	2	>.05
Avoid_emb S2	1.98			
Avoid_emb S3	1.96			
Improve_grade S1	2.28	6.035	2	<.05
Improve_grade S2	1.76			
Improve_grade S3	1.96			
LB_is_fun S1	2.22	2.844	2	>.05
LB_is_fun S2	1.90			
LB_is_fun S3	1.88			
Activities_neg S1	1.98	1.069	2	>.05
Activities_neg S2	1.90			
Activities_neg S3	2.12			
Motivated_Eng S1	2.20	2.471	2	>.05
Motivated_Eng S2	1.96			
Motivated_Eng S3	1.84			
Remind_Eng S1	2.22	3.265	2	>.05
Remind_Eng S2	1.80			
Remind_Eng S3	1.98			

Note. ** = $p < .05$

4.3.7 Measuring Attitudes Based on Leaderboard Rank: ANOVA Results

One-way analysis of variance (ANOVA) tests were conducted on the 12 statements to determine whether there were significant differences in attitudes towards the leaderboards depending on a participant's final leaderboard rank. The independent variable was final leaderboard rank (high, middle, low), and the dependent variable was the participants' mean scores towards each statement. Further details about how the participants were ranked based on final leaderboard ranking are shown in Section 4.4.4.

Prior to the analysis, one-way between groups ANOVA assumptions were examined (Field, 2013). The assumptions of univariate normality of residuals, homoscedasticity of residuals, and the lack of outliers were assessed. Normality was evaluated using Q-Q scatterplots that compare the distribution of the residuals with a normal distribution. Normality was assumed if the points form a relatively straight line. Homoscedasticity was evaluated by plotting the residuals against the predicted values. The assumption of homoscedasticity was met if the points appeared randomly distributed with a mean of zero and no apparent curvature. To identify influential points, studentized residuals were calculated and the absolute values were plotted against the observation numbers. Studentized residuals were calculated by dividing the model residuals by the estimated residual standard deviation. An observation with a studentized residual greater than 3.45 in absolute value, or the .999 quartile of a *t* distribution with 25 degrees of freedom, was considered to have significant influence on the results of the model. No assumptions were violated.

Table 4.12 presents the descriptive statistics for each statement. Table 4.13 displays the results of the one-way ANOVAs. The results of ANOVA tests were significant at the $p < .05$ level for three statements: Maintain_rank, LB_is_fun, and Activities_neg. Post Hoc analyses using paired samples *t*-tests were conducted on the statistically significant results. Tukey pairwise comparisons were conducted for all significant effects.

Table 4.12
Descriptive Statistics for Leaderboard Statements Based on Final Rank

statement	Final rank	*n*	*M*	*SD*
Gain_rank	high	9	4.56	1.04
	middle	9	4.67	1.03
	low	8	4.33	.62
Maintain_rank	high	9	5.22	.58
	middle	9	4.96	.82
	low	8	3.79	1.08
Don’t_care	high	9	2.26	1.23
	middle	9	2.85	1.31
	low	8	3.67	1.23
Show_teacher	high	9	4.19	1.11
	middle	9	4.26	.81
	low	8	3.46	.91
Show_students	high	9	3.33	1.56
	middle	9	3.74	1.15
	low	8	2.96	.88
Prove_to_myself	high	9	4.56	1.08
	middle	9	4.11	1.09
	low	8	3.46	.89
Avoid_emb	high	9	4.33	1.27
	middle	9	4.00	1.22
	low	8	3.88	1.27
Improve_grade	high	9	4.70	1.20
	middle	9	4.56	1.05
	low	8	4.46	.80
LB_is_fun	high	9	4.67	.82
	middle	9	4.35	.90
	low	8	3.42	1.12
Activities_neg	high	9	2.11	.90
	middle	9	2.35	.82
	low	8	3.38	1.37
Motivated_Eng	high	9	4.85	.71
	middle	9	4.37	1.12
	low	8	3.62	1.16
Remind_Eng	high	9	4.85	.82
	middle	9	4.11	1.20
	low	8	3.54	1.26

Table 4.13
Merged Analysis of Variance Tables for All Leaderboard Statements

		Sum of Squares	*df*	*F*	Sig.	η^2
Gain_rank	Between	0.48	2	.28	.758	.02
	Residuals	19.78	23			
Maintain_rank	Between	9.68	2	6.82	.005*	.37
	Residuals	16.31	23			
Don't_care	Between	8.42	2	2.64	.093	.19
	Residuals	36.64	23			
Show_teacher	Between	3.26	2	1.80	.188	.14
	Residuals	20.85	23			
Show_students	Between	2.60	2	.84	.444	.07
	Residuals	35.60	23			
Prove_to_myself	Between	5.13	2	2.41	.112	.17
	Residuals	24.43	23			
Avoid_emb	Between	0.97	2	.31	.738	.03
	Residuals	36.21	23			
Improve_grade	Between	0.26	2	.12	.886	.01
	Residuals	24.75	23			
LB_is_fun	Between	7.06	2	3.93	.034*	.25
	Residuals	20.64	23			
Activities_neg	Between	7.50	2	3.45	.049*	.23
	Residuals	25.01	23			
Motivated_Eng	Between	6.43	2	3.14	.062	.21
	Residuals	23.55	23			
Remind_Eng	Between	7.36	2	3.02	.068	.21
	Residuals	28.01	23			

** = $p < .05$

For Maintain rank, the ANOVA results $F(2, 23) = 6.82$, $p = .005$ indicate a significant difference among the levels of Final rank. The eta squared was .37 indicating Final rank explains approximately 37% of the variance in Maintain rank. Post hoc *t*-test results showed significant difference for two pairs.

The mean score for the Final rank high group ($M = 5.22$, $SD = .58$) was significantly larger than the low group ($M = 3.79$, $SD = 1.08$), $p = .005$), and the mean score for

the Maintain rank middle group (M = 4.96, SD = .82) was also significantly larger than the low group (M = 3.79, SD = 1.08), p = .023). Figure 4.3 visually presents this data.

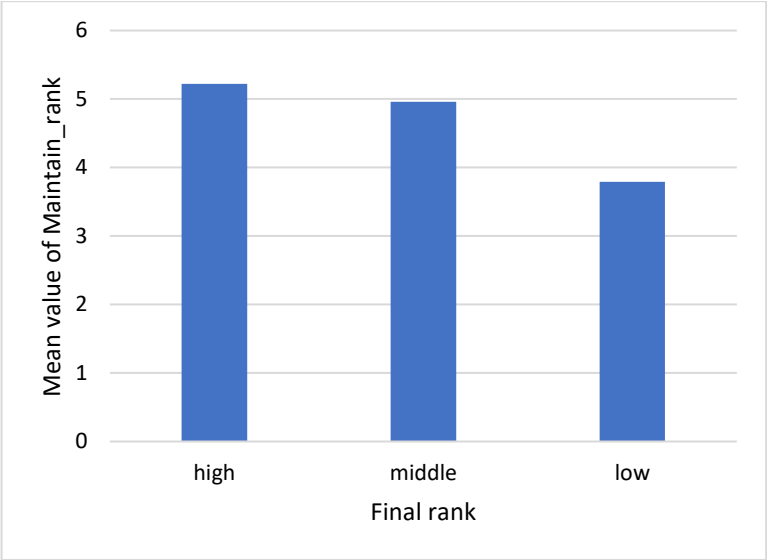


Figure 4.3. Maintain rank means by factor levels of Final rank.

For LB_is_fun, the ANOVA results $F(2, 23) = 3.93$, p = .034 indicate a significant difference in LB_is_fun among the levels of Final rank. The eta squared was .25 indicating Final rank explains approximately 25% of the variance in LB_is_fun. Post hoc t-test results showed a significant difference for one pair. The mean score for the LB fun high group (M = 4.67, SD = .82) was significantly larger than for the low group (M = 3.42, SD = 1.12), p = .032. Figure 4.4 visually presents this data.

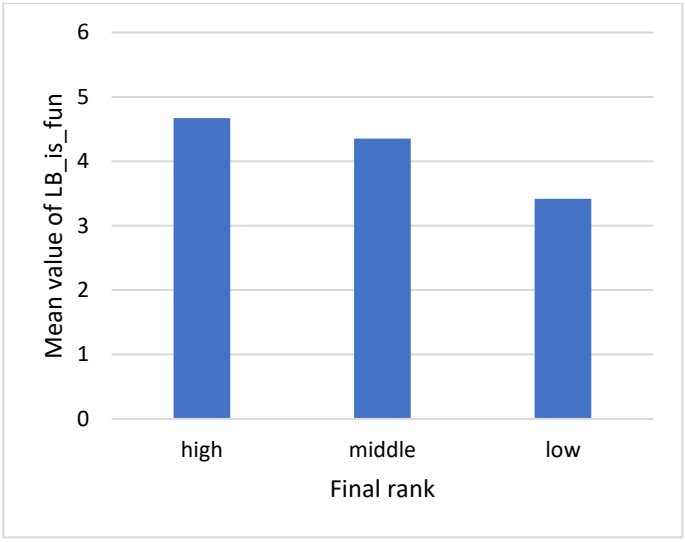


Figure 4.4. LB_is_fun means by factor levels of Final rank.

For Activities_neg, the ANOVA results $F(2, 23) = 3.45$, $p = .049$, indicate a significant difference in Activities_neg among the levels of Final rank. The eta squared was .23 indicating Final rank explains approximately 23% of the variance in Activities_neg. Post hoc *t*-test results showed no significant difference between the pairs. However, the data presented in Figure 4.5 suggest that the lower ranked participants had stronger negative feelings towards the learning activities compared to the middle and high groups.

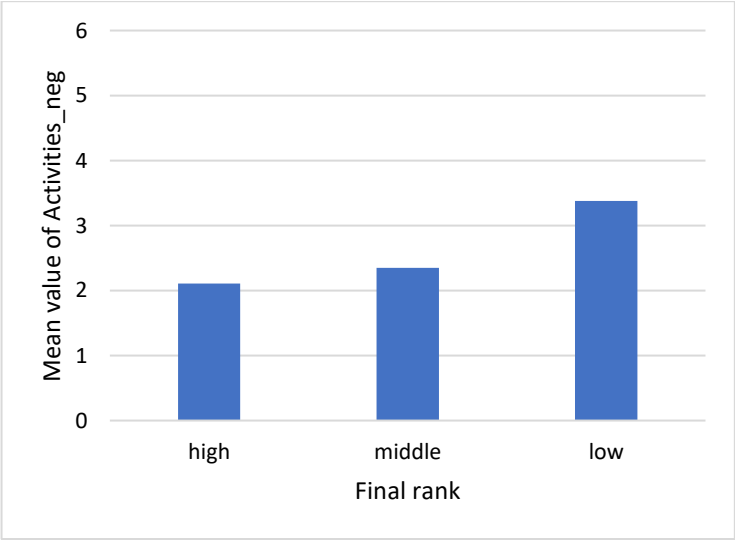


Figure 4.5. Activities_neg means by factor levels of Final rank.

4.3.8 Leaderboard Questionnaire: Opinion Section Content Analysis Results

The final section of the leaderboard questionnaire asked the participants to respond to the non-mandatory open-ended question "Generally, what do you think about using leaderboards in class?" After the completion of Survey 3, a quantitative thematic content analysis was conducted on the responses. Responding to this question was non-mandatory; therefore, the data only reflect the opinions of the participants that wrote a response. The high number of participants that did not write a response could reflect some of the participants' survey fatigue after already answering many questions about leaderboards in the previous sections of the leaderboard questionnaire. Table 4.14 presents the results of the analysis.

The results of the content analysis were positive overall towards leaderboards. The three most frequently coded thematic categories were: (1) Good for motivation, (2) It's good, (3) Points are good. An example entry coded into the Good for motivation category was Participant 4's Survey 1 response: "I think it encourages us to work hard." An example entry coded into the Points are good category was Participant 7's Survey 2 response: "I think using Leaderboards are good for us because we can check our score and see if we need to work harder or not." The other thematic categories are self-explanatory based on their thematic category titles in Table 4.14.

Table 4.14
Opinions towards Leaderboards Content Analysis Results

Thematic category	Survey 1 *n = 26*	Survey 2 *n = 25*	Survey 3 *n = 26*	Total	Rank
Good for motivation	8	4	3	15	1
It's good	3	3	5	11	2
Points are good	3	4	0	7	3
I don't mind	1	1	1	3	4
Don't like	0	1	1	2	5
No answer given	11	12	16	39	
Total	26	25	26	77	

4.4 Quest Diary Results

All the participants (N = 46) wrote in their quest diaries six times, approximately every two weeks, coinciding with the completion of a quest level. At the end of the quest diary data collection period, a total of 240 quest diaries were received; 41 for Level 2, 38 for Level 3, 40 for Level 4, 39 for Level 5, 39 for Level 6, and 43 for the final quest diary. Class 1 (n = 26) and Class 2 (n = 20) completed 120 quest diaries each. 140 entries were written in Japanese, 95 were written in English, and 5 were written in a mixture of Japanese and English. The Japanese entries were translated into English by a professional translator in preparation for the data analysis.

Guided by the literature (Berg & Lune, 2012; Cohen et al., 2011; Hatch, 2012), a thematic quantitative content analysis was conducted on the quest diaries using NVivo 12. NVivo is a computer program used for qualitative data analysis in which the user codes segments of data to user-created nodes that represent areas of research interest for the user. The data were iteratively coded using a variety of recommended coding schemes (Auerbach & Silverstein, 2003; Miles & Huberman, 1994; Namey, Guest, Thairu, & Johnson, 2008) that were suitable for this study. Descriptive codes tracked demographic data such as participant number, class number, and gender. Structural codes labelled specific data such as answers to specific questions.

Open coding was conducted as an exploratory process during the initial readings of the texts by attaching meaningful descriptive labels to areas of text that were of interest to the researcher; multiple open codes could have been applied to the same piece of text. Sentiment codes were attached to the open codes to delineate between positive or negative comments. Thematic axial coding was conducted to categorise open codes into emerging thematic categories. The open codes and axial codes were developed inductively from the emerging themes in the text, and deductively based on the goals of this study. As recommended by Lynch (2003), an external code checker checked the accuracy of the codes by coding a section of the text based on the codebook. The inter-coder reliability results were acceptable based on the minimal benchmark of 85-90% (Saldaña, 2009).

The results of the quest diary content analyses are separated into seven topics: two

topics are about quest choice, three topics are about opinions towards the quests, and two topics are about opinions towards the class. The topics represent seven different questions that the participants responded to. For each topic, the frequency of each node was calculated and then ranked based on their frequency. The data is presented in tables that present only the highest frequency and most pertinent nodes and categories. For each table, *n* refers to the number of different participants whose response corresponded with that node at least once, and *freq* refers to the frequency of a node being coded. It is important to distinguish between *n* and *freq* because one participant's response could have been coded to multiple nodes, multiple times, over the duration of the six quest diaries. The rank refers to that node's ranking compared to the other nodes. The frequency and ranking of each node suggest its importance for understanding the participants' perceptions towards quests.

4.4.1 Opinions of Individual Quests Content Analysis Results

At the end of each level, the participants wrote their opinions about the quests they recently completed. Table 4.15 shows the results of the content analysis that was performed on the participants' responses. The results do not represent any single quest but do represent a combination of all the responses to each completed quest. Two main categories of response evolved from the data are learning and personal feelings. The learning category subsumes seven nodes that relate to the participants responding that the quest they completed led to a positive education-related learning outcome; 38 participants (95%) were coded into the learning category 296 times (32% of total codes). The personal feelings category subsumes six nodes that reflect the participants' personal feelings towards the quests they completed; 40 participants (100%) were coded into the personal feelings category 608 times (67% of total codes). Figure 4.6 provides a visual representation of the data reflecting the participants' opinions towards the individual quests.

Table 4.15
Opinions towards Individual Quests Content Analysis Results

Question: What did you think about the quest you completed?	*n* (40)	*n* %	*n* *rank*	*freq* (909)	*freq* %	*freq* rank
I. Learning	38	95		296	33	
i. Improved or practiced English	37	93	2	136	15	2
ii. International knowledge	27	68	7	54	6	6
iii. New way to learn English	28	70	6	46	5	8
iv. I.T. skills	22	55	9	32	4	9
v. Used English to learn something	9	23	11	10	1	12
vi. General knowledge	6	15	12	11	1	11
vii. Mobile learning	5	13	13	7	1	13
II. Personal feelings	40	100		608	67	
i. General positive feeling	38	95	1	222	24	1
ii. Made me reflect	36	90	3	113	12	3
iii. Motivation	34	85	4	102	11	4
iv. Challenging in some way but could succeed	33	83	5	91	10	5
v. Interpersonally positive	26	65	8	49	5	7
vi. Negative feelings	17	43	10	31	3	10

Note. *f* = node citation frequency

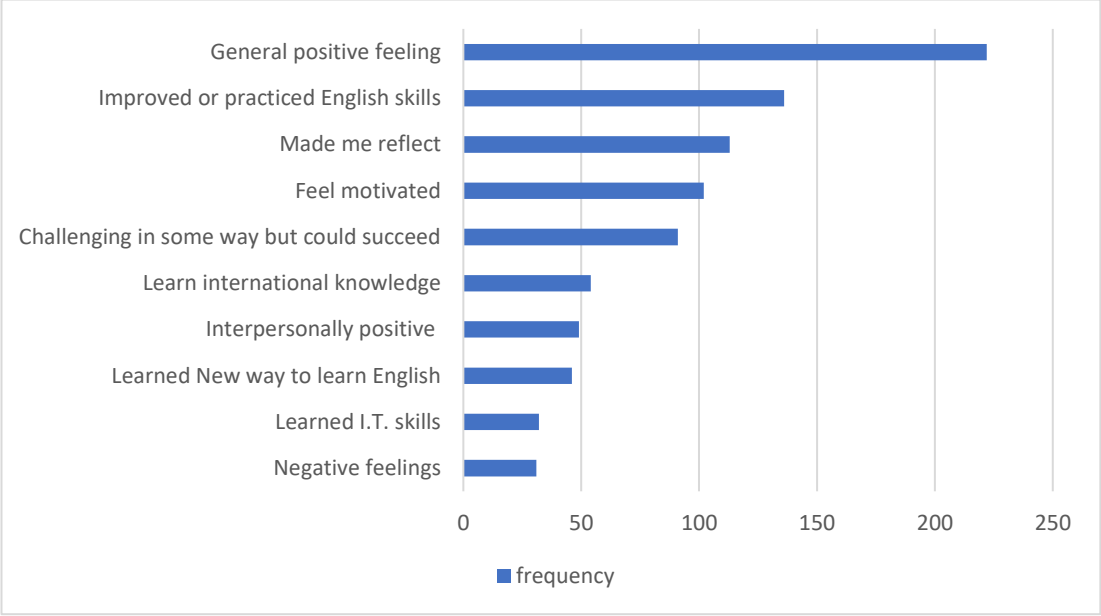


Figure 4.6. Top 10 most frequent opinions about individual quests.

The most frequently coded node was *general positive feelings*; 38 participants (95%) were coded to the node 222 times (24% of total codes). The sub nodes of the *general positive feelings* node show that 32 participants (80%) said that a *quest was meaningful and/or useful* 123 times (14% of total codes), 33 participants (82%) said that a *quest was enjoyable* 75 times (8% of total codes), 11 participants (27%) said they *felt sense of achievement when quest was finished* 14 times (2% of total codes), 5 participants (12%) said they were *happy to be able to choose learning content* six times (.5 % of total codes), and 4 participants (10%) said that a quest *didn't feel like homework* four times (.5% of total codes).

The second most frequently coded node was *improved or practiced English skills*; 37 participants (93%) were coded to the node 136 times (15% of total codes). The node comprises 6 sub nodes which show the specific skills the participants referred to: 23 participants (60%) mentioned *listening* a total of 45 times (5% of total codes); 22 participants (55%) mentioned *speaking* 30 times (3% of total codes); 18 participants (45%) mentioned *writing* 27 times (3% of total codes); 12 participants (30%) mentioned *presentation skills* 15 times (2% of total codes); 8 participants (20%) mentioned *pronunciation* 10 times (1% of total codes); and 8 participants (20%) mentioned *vocabulary* 8 times (1% of total codes).

The third most frequently coded node was *made me reflect*; 36 participants (90%) were coded to the node 113 times (12% of total codes). The sub nodes show what the participants reflected on: 28 participants (70%) said *themselves* 36 times (4% of total codes); 15 participants (38%) said their *language skills* 19 times (2% of total codes); 18 participants (45%) said *the people around them who have helped them* 18 times (2% of total codes); 13 participants (33%) said their *major* 13 times (1%); 11 participants (28%) said their future travel plans 12 times (1% of total codes); 8 participants (20%) said their future work 9 times (1% of total codes); and 5 participants (13%) said social issues 6 times (.5% of total codes).

The fourth most frequently coded node reflected a feeling of *motivation*; 34 participants (85%) were coded to the node 102 times (11% of total codes). The sub nodes show what the participants were motivated to do: 23 participants (58%) said *improve English skills* 44 times (5% of total codes); 20 participants (50%) said *do*

positive things for their own lives 27 times (3% of total codes); 12 participants (30%) said *do positive things for others* 12 times (1% of total codes); 11 participants (37%) said *get international exposure* 12 times (1% of total codes); and 7 participants (18%) said *improve their presentation skills* 7 times (.5% of total codes).

The fifth most frequently coded node was *challenging but could be completed successfully in the end*; 33 participants (83%) were coded to the node 91 times (10% of total codes). The sub nodes show what aspect of a quest a participant found challenging: 26 participants (65%) said their *language ability* 54 times (6% of total codes); 17 participants (43%) said their *I.T. skills* 19 times (2% of total codes); 12 participants (30%) said to *reflect on themselves* 12 times (1% of total codes). The sixth most frequently coded node represents the participants' comments about the *international knowledge* they learned; 27 participants (68%) were coded to the node 54 times (6% of total codes). The sub nodes show what aspects of international culture the participants learned about: 24 participants (60%) said *foreign countries* 33 times (4% of total codes); 8 participants (20%) said *culture* 11 times (1% of total codes); and 9 participants (23%) said *different ways of thinking* 9 times (1% of total codes).

The seventh most frequently coded node represents the participants' comments about a quest providing an *interpersonally positive* experience; 26 participants (65%) were coded to the node 49 times (5% of total codes). The sub nodes show what aspects of a quest were interpersonally positive: 15 participants (38%) said *enjoyed sharing their life with other classmates* 24 times (3% of total codes); 8 participants (20%) said *enjoyed working with other people to complete a quest* 10 times (1% of total codes); 4 participants (10%) said *enjoyed learning about their classmates* 11 times (1% of total codes); and 5 participants (13%) said *sharing was motivating*, specifically, sharing personal goals or seeing how other participants completed the quest made them want to improve their own completed quest six times (.5% of total codes).

The eighth most frequently coded node, *new way to learn English*, represents the participants' positive comments about learning a new way to study English that they will continue to use in the future; 28 participants (70%) were coded to the node 46

times (5% of total codes). The 13th most frequently coded node, mobile learning, represents the participants' positive comments from the *new way to learn English* sub node about being able to learn while on the train using their smart phone; 5 participants (13%) were coded to the node 7 times (.5% of total codes).

The ninth most frequently coded node, *I.T. skills*, represents the participants' positive comments about improving their I.T. skills, specifically, learning how to make narrated PowerPoint presentations and videos, using websites to share their completed work; 22 participants (55%) were coded to the node 32 times (4% of total codes). The 11th most frequently coded node, *general knowledge*, represents the participants' comments about learning about Japan and environmental issues such as climate change; 6 participants (15%) were coded to the node 11 times (1% of total codes). The 12th most frequently coded node, *used English to learn something*, represents the participants' positive comments about a quest allowing them to use English in a meaningful way to learn something that was non-language related; 9 participants (23%) were coded to the node 10 times (1% of total codes).

The tenth most frequently coded node, *negative feelings*, represents the participants' negative opinions towards a quest they completed; 17 participants (43%) were coded to the node 31 times (3% of total codes). The sub nodes show what aspects of the quest a participant had a negative opinion towards: 11 participants (28%) said a quest was *difficult* 18 times (2% of total codes); 7 participants (18%) said a quest was *time consuming* 9 times (1% of total codes); and 4 participants (10%) said they *did not want to speak in front of a camera* 4 times (.5% of total codes).

4.4.2 Perceptions of Quest-based Learning Content Analysis Results

In the final quest diary, the participants were asked to write about what they liked and did not like about quests as an approach to learning English. Thematic quantitative content analysis was conducted on the responses. Each unique theme in each response was coded. One response could have been coded for multiple unique themes. 96 themes were coded for the 42 responses. Each response was also coded for overall sentiment. If over half of a response was positive, then it was coded as positive; if it was over 50% negative it was coded as negative; and if the response

was half positive and half negative, it was coded as neutral. For example, one participant's response was coded for sentiment as this: "Sometimes I felt it's hard and tiresome <negative>, I knew I could learn a new skill if I tried <positive>. I enjoyed writing in my blog and reading other students' blogs <positive>". The overall sentiment of this response was coded as positive as the response had over 50% positive codes. The frequency of the sentiment codes assigned to the 42 responses were: positive (n = 34), neutral (n = 6), and negative (n = 2). The frequency of the positive sentiment codes suggests that the participants generally had stronger positive opinions towards quests than negative opinions towards quests.

Table 4.16 displays the results of the content analysis. It shows that many of the participants wrote a mixture of positive and negative comments. As the participants were told to write about the things they liked and disliked about quests, the results do not represent overall sentiment but highlight areas of interest. Figure 4.7 provides a visual representation of the participants' perceptions of the QBL experience.

Table 4.16

Participants' Perceptions of QBL for Learning EFL Content Analysis Results

Question: What do you like or dislike about quests?		*n (n = 42)*	rank
I.	Positive	40	
i.	Good for English education	17	1
ii.	Appreciated social aspect	11	3
iii.	Provided meaningful opportunities	9	5
iv.	Choice of various quests was good	8	6
v.	Weekly flow of quests was good	5	7
vi.	Challenging but could learn	5	7
vii.	Enjoyed quests	4	9
viii.	Felt motivated by completing quests	2	10
ix.	Could improve I.T. skills	2	10
II.	Negative	28	
i.	Technically difficult	12	2
ii.	Time consuming	11	3
iii.	A little stressful when busy	2	10
iv.	Not all quests were interesting and motivating	2	10

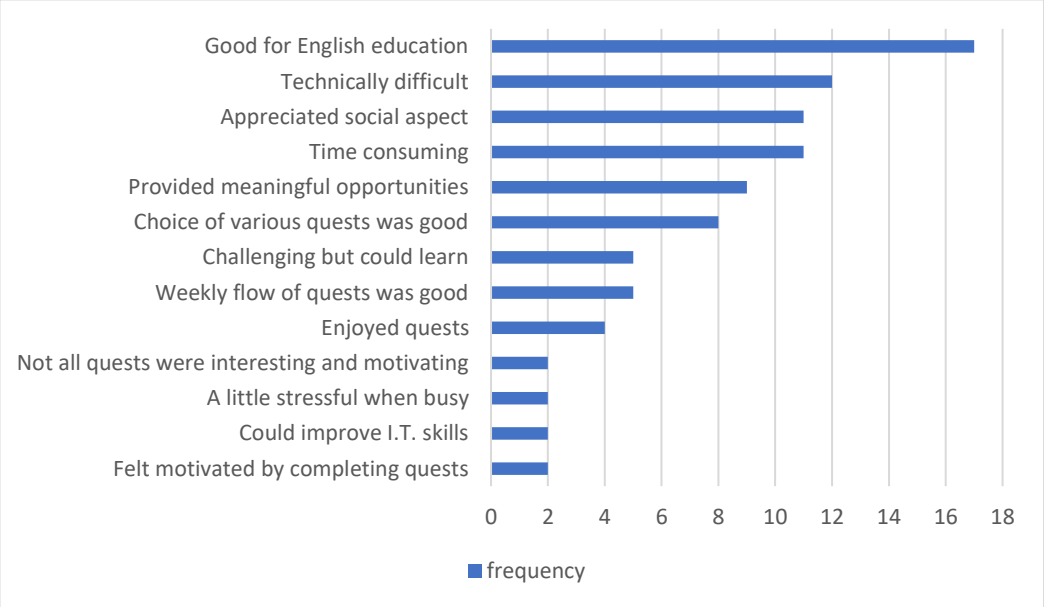


Figure 4.7. Participants' general perceptions towards the QBL implementation.

Table 4.16 shows the variety of positive perceptions the participants had towards the quests and the QBL implementation. The most frequently coded node overall was *good for English education*. From the 17 participants (40%) who were coded to the node, 9 participants (21%) wrote specifically that through the quests they could use or improve their English skills, 4 participants (10%) wrote that they learned many news ways to study English, and 4 participants (10%) wrote positively about the quests allowing them to improve their English skills by using English in a meaningful situations. The third most frequently coded node was *appreciated social aspect*. The 11 participants (26%) who were coded to the node explained that they liked learning about and from other classmates, and liked writing and sharing weekly blog. The other positive nodes show that 9 participants (21%) commented that quests provided them with meaningful opportunities to experience something, 8 participants (19%) commented that a choice of quest was good, 5 participants (12%) commented that the weekly flow of quests was good, 5 participants (12%) commented that the quests were challenging, but, in the end, they learned something, 4 participants (10%) commented that they just generally enjoyed doing the quests, 2 participants (5%) commented that they felt motivated by completing quests, and 2 participants (5%) commented that they improve their I.T. skills.

Table 4.16 also shows that, compared to the wide range of positive perceptions the participants had towards the quests, the variation of negative perceptions was much less. The second most frequently code node overall was *technically difficult*; the 12 participants (29%) coded to this node often explained that the process of making narrated PowerPoint presentation videos and publishing them on their blogs was difficult because they were not good at using computers. The third most frequently coded node was *time consuming*; the 11 participants (26%) coded to this node explained that some of the quests took a long time to complete. Finally, 2 participants (5%) commented that quests were *a little stressful when they were busy*, and 2 participants (5%) commented that *not all the quests were interesting and motivating*.

4.4.3 Comparing Attitudes towards the Learning Activities: Descriptive Statistics

In the final quest diary, the participants were asked whether they thought the learning activities used in the class were good methods for improving their English ability. For each learning activity they responded to a 6-point Likert scale anchored by 1 (*strongly disagree*) and 6 (*strongly agree*). Table 4.17 presents the mean scores and rankings for each learning activity for each class. The results show positive attitudes towards all the learning activities; however, the mean scores for the textbook are noticeably lower than the three activities that made up the gamified portion of the class. Therefore, the data suggest that the students preferred the quests, English Central, and extensive reading more than the textbook.

Table 4.17
Participants' Opinions towards the Learning Activities

	Class 1 (*n = 22*)			Class 2 (*n = 20*)			Class 1&2 combined (*n = 42*)		
	M	*SD*	rank	*M*	*SD*	rank	*M*	*SD*	rank
quests	5.05	0.95	2	5.40	0.68	1	5.22	0.84	1
English Central	5.27	0.94	1	5.15	0.93	2	5.21	0.92	2
extensive reading	5.00	0.98	3	5.05	1.28	3	5.03	1.12	3
textbook	4.41	1.14	4	4.35	1.09	4	4.38	1.10	4

4.4.4 Opinions of the Class: Content Analysis Results

In the final quest diary survey, all the participants were asked to write their general opinion about the class. A thematic quantitative content analysis was performed on the responses. Each unique theme in each response was coded. One response could have been coded for multiple themes. Table 4.18 presents the most frequent nodes and their associated rankings. A matrix query was run in NVivo 12 to determine if there were any major differences in the responses between Class 1 and Class 2. No

major differences were found.

Table 4.18
Participants' General Opinions towards the Class

Question: What did you think about this class this semester?		*n* (*n* = 43)	rank
i.	Authentic and meaningful English usage and improvement	21	1
ii.	Enjoyment, fun, good, interesting	21	1
iii.	Challenging but worthwhile	19	3
iv.	Class structure	13	4
v.	New experiences	12	5
vi.	Could make friends and develop relationships	9	6
vii.	Feel motivated	6	7
viii.	Nice teacher	5	8
ix.	Heavy workload	2	9

Table 4.18 shows that 21 participants (49%) said that the class allowed them to improve and use their English skills in authentic and meaningful situations, 21 participants (49%) commented that the class was enjoyable, fun, good, or interesting, and 19 participants (44%) mentioned that the amount of work was challenging but it was worthwhile. The table shows that 13 participants (30%) wrote about the class structure: 6 participants (14%) said that it took time to get used to the class; 3 participants (7%) said it was good to learn English in different ways; 2 participants (5%) said it was different than other classes; and 2 participants (5%) said that they liked this style of class. The table shows that 12 participants (28%) commented that they had positive new experiences such as doing quests, writing blogs, and making narrated PowerPoint presentations. The participants also wrote about the interpersonal aspect of the class: 9 participants (21%) said that they could make friends and develop relationships, and 5 participants (12%) said they liked the teacher. The table shows that that 6 participants (14%) wrote that the class made them motivated to study; the reasons for motivation include the weekly point scoring system, the leaderboard, and the quests. Finally, 2 participants (5%) wrote negatively about the workload being too heavy.

4.4.5 Reasons why a Quest was Chosen Content Analysis Results

At the end of each level, the participants wrote about why they chose to do a specific quest. Table 4.19 shows the results of the content analysis that was performed on the participants' responses. Four main categories of response evolved from the data: personal experience, learning, content appearance, and avoidance. Each category subsumes multiple related sub nodes. Figure 4.8 visually presents the ranked *freq* data.

Table 4.19
Reasons Explaining Why a Participant Chose to do a Quest Content Analysis Results

Question: Why did you choose to do that specific quest?	*n* (40)	*n* %	*n* rank	*freq* (355)	*freq* %	*freq* rank
I. Personal experience	33	82		125	35	
i. Wanted to reflect on my life	23	57	1	31	9	4
ii. Wanted to do this type of quest	18	45	6	23	6	6
iii. Wanted to share something about myself	12	30	11	19	5	8
iv. Wanted to achieve a goal	17	42	8	18	5	9
v. Wanted to try something new	12	30	11	16	5	10
vi. Seemed worthwhile or useful	8	20	15	10	3	14
vii. Wanted to work with my friend	5	12	7	8	2	16
II. Learning	32	80		101	28	
i. Wanted to improve or use English skill	23	57	1	46	13	2
ii. Wanted to learn about foreign countries	20	50	5	29	8	5
iii. Wanted to improve computer skills	13	32	9	15	4	11
iv. Wanted to learn about my major	11	27	13	11	3	13
III. Content appearance	30	75		83	23	
i. Looked easy or quick to finish	18	45	6	47	13	1
ii. Looked interesting, fun, enjoyable	22	55	3	36	10	3
IV. Avoidance	21	52	4	46	13	
i. Making a video is problematic	13	32	9	23	6	6
ii. Didn't want to work with other people	9	22	14	14	4	12
iii. Difficult for some reason	6	15	16	9	3	15

Note. *n* = number of participants who responded at least once, *freq* = node citation frequency.

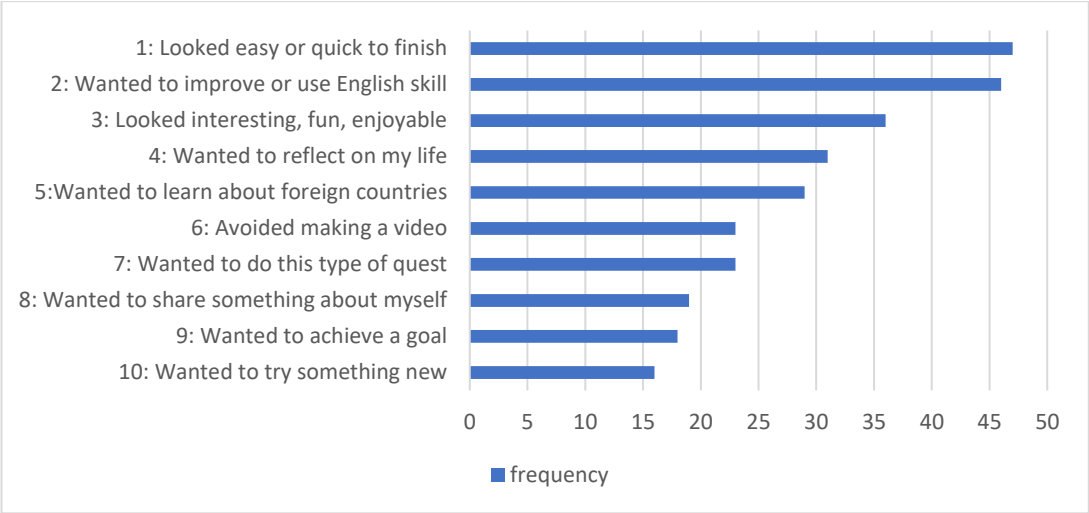


Figure 4.8. Top 10 reasons why a quest was chosen.

The personal experience category represents nodes that relate to the participants choosing to do a quest because they wanted to experience something in their life. The personal experience category was the most common category of response with 33 participants (82%) coded into it at least once over the six quest diary entries. Seven nodes, ranked from the most common to the least common, represent the different types of experiences the participants said they wanted to have: 23 participants (57%) were coded to the node *wanted to reflect on my life* 31 times; 18 participants (45%) were coded to the node *wanted to do this type of quest* 23 times; 12 participants (30%) were coded to the node *wanted to share something about myself* 19 times; 17 participants (42%) were coded to the node *wanted to achieve a goal* 18 times; 12 participants (30%) were coded to the node *wanted to try something new* 16 times; 8 participants (20%) were coded to the node *seemed worthwhile or useful* 10 times; and 5 participants (12%) were coded to the node *wanted to work with a friend* 8 times. The *wanted to reflect on my life* node was the shared number one *n* ranked code overall; however, it was only the fourth most frequently coded node overall.

The learning category represents nodes that relate to the participants choosing to do a quest because they wanted to learn about something or improve an education-related skill. The learning category was the second most common category of response with 32 participants (80%) coded into it at least once over the six quest diary entries. Four nodes, ranked from the most common to the least common, represent the different areas the participants said they wanted to learn or practice: 23 participants (57%)

were coded to the node *wanted to improve or use English skill* 46 times; 20 participants (50%) were coded to the node *wanted to learn about foreign countries and culture* 29 times; 13 participants (32%) were coded to the node *wanted to improve computer skills* 15 times; 11 participants (27%) were coded to the node *wanted to learn about my major* 11 times. The *wanted to improve or use English skill* was the shared number one *n* ranked code and it was also the second most frequently coded node overall.

The content appearance category represents nodes that relate to the participants choosing to do a quest based on the superficial appearance of the quest. The content appearance category was the third most common category with 30 participants (75%) being coded into it at least once over the six quest diary entries. The sub nodes show that 18 participants (45%) were coded to the node *looked easy or quick to finish* 47 times, and 22 participants (55%) were coded to the node *looked interesting, fun, enjoyable* 36 times. The *looked easy or quick to finish* node was the most frequently coded node overall with 47 instances of it being recorded; however, as the instances only came from 18 participants (45%), it was only the sixth *n* ranked code overall.

The avoidance category represents nodes that relate to the participants choosing to not do a quest for some reason. 21 participants (52%) were coded into the avoidance category at least once over the six quest diary entries. The avoidance category subsumes three sub nodes. 13 participants (32%) were coded to the node *making a video is problematic* 23 times; 8 participants (20%) said that it was difficult due to time constraints, 3 participants (7%) said that they felt too embarrassed to appear on a video, and 4 participants (10%) said that they just did not want to make a video. 9 participants (22%) were coded to the node *didn't want to work with other people* 14 times; four participants (10%) said that it was due to time constraints, four participants (10%) said that they did not have any foreign friends to work with, two participants (5%) said that it was troublesome to work with classmates, and two participants (5%) said they just wanted to work by themselves. Six participants (15%) were coded to the node *difficult for some reason* nine times; four participants (10%) explained that they did not have the required content to complete the quest.

4.4.6 Quest Choice: Who Should Decide what Quest to Work on

In the final quest diary, the participants (n = 42) shared their opinions about whether the teacher or the participant should choose which quest to do. Table 4.20 shows the results of the content analysis that was performed on the participants' responses. The data clearly show that most of the participants prefer choosing for one of three main reasons: (1) they can choose quests they are interested in; (2) they can choose quests depending on their workload or life; or (3) they feel motivated when they can choose what quest to work on.

Table 4.20
Who Should Choose the Quest Content Analysis Results

Question: Who should choose what quest to do and why?	*n* *42*	rank
I. Student	39	
i. Can choose quests I'm interested in and enjoy more	15	1
ii. Can choose quests depending on workload or life	9	2
iii. Feel motivated if I have freedom of choice	9	2
iv. Can choose based on difficulty or quest	6	4
II. Teacher	3	
i. Hard to decide	2	5

4.4.7 Opinions of the Collaboration Aspect of Quest-based Learning Results

In the final quest diary, the participants were asked to write their opinion about the collaboration aspect of quests. The responses (n = 41) were analysed using a quantitative content analysis. Three major categories of response evolved from the data: "positive", "negative", and "did not collaborate". Table 4.21 presents the results of the content analysis. The results show that, out of the 41 participants who provided a response, 28 (68%) were positive, 2 (5%) were negative, and 11 (27%) participants said that they did not collaborate. For the positive responses, 12 participants (29%) said that the quests were a good opportunity to build relationships and make friends, 7 participants (17%) said that collaborating was enjoyable, and 6

participants (15%) said that the collaboration aspect of questing provided them with good opportunities to interact with members of the class. For the participants who said that they did not collaborate, 5 participants (12%) said that they chose quests they could do by themselves, 4 participants (10%) said that they were too busy or shy to work with people, and 2 participants (5%) said that they could not schedule the time to work with other people.

Table 4.21
Participants' Responses about the Collaboration Aspect of Questing

Question: What did you think of the collaboration aspect of questing?	*n* *(41)*	*n* %	rank
I. Positive	28	68	1
i. Good opportunity to build relationship make	12	29	
ii. Fun, good, or enjoyable	7	17	
iii. Good opportunity to interact with classmates	6	15	
II. Did not collaborate	11	27	2
i. I chose quests I could do by myself	5	12	
ii. Too busy, too shy, or avoided video quests	4	10	
iii. Couldn't schedule with other classmates	2	5	
III. Negative	2	5	3
i. Did not like it	1	2	
ii. I am not good at working with other people	1	2	

4.5 Semi-structured Interview Content Analysis Results

Semi-structured interviews were conducted with 9 participants from Class 1 soon after the Intensive English course had concluded. The audio recordings of the interviews were transcribed into Japanese and then translated into English. A thematic content analysis was conducted on the English transcriptions using NVivo 12. This section presents the results of the content analysis in relation to the participants' comments about quests and leaderboards. A comparison between the participants who had a high final leaderboard ranking (n = 5) to those who had a middle or low final leaderboard ranking (n = 4) is presented to compare data based

on leaderboard rank. The results presented in this section provide the reader with a quick general summary of the semi-structured interview findings. Individual interview summaries can be found in Appendix F.

Table 4.22 presents the results of the content analysis in relation to quests. The results are categorised into relevant themes that emerged from the data. The themes are axial codes that subsume several positive, neutral, and negative sub nodes about the theme. The greater frequency of positive nodes suggests that the participants generally felt positively towards the quests. All nine participants commented that being able to choose which quest to work on was good. Five participants commented that the quests were initially difficult but became easier as they became more comfortable with the questing process; this comment was more prevalent in the high group (n = 4) than the middle and low group (n = 1. Five participants said they enjoyed the social aspect of questing; however, two participants said they preferred quests they could complete by themselves. More detailed accounts of the participants' perceptions towards quests are presented in Sections 5.3 and 5.4.

Table 4.22
Summary of Interview Responses about Quests

Theme		Response	Final leaderboard rank		total
			High (n = 5)	mid & low (n = 4)	(n = 9)
Choosing quests	+	Being able to choose quests was good	5	4	9
	~	Did not choose video making quests	2	0	2
Opinions about quests	+	Initially difficult but became easier	4	1	5
	+	Good way to study English	3	0	3
	+	Appreciated reflecting on life	2	0	2
	+	Enjoyed the creative aspect	1	1	2
	+	Learned new ways to Study English	3	0	3
	-	Did not like quests that took a long time	1	2	3
Collaboration	+	Enjoyed the social aspect	2	3	5
	-	Preferred single person quests	1	1	2
Technical design of quests	+	Quest frequency was good	1	0	1
	~	Wanted time in class to do quests	0	1	1
	-	Time discrepancy between quests existed	0	1	1

Note. + = positive comment, ~ = neutral comment, - = negative comment.

Table 4.23 presents the results of the content analysis in relation to leaderboards. The results are categorised based on their valence and frequency. The greater frequency of positive nodes suggests that the participants generally felt positively towards the leaderboards. The results show that the high group had more positive comments about the leaderboards, and the middle and low groups had more negative comments about the leaderboards. For example, from the high group, 3 participants commented that the leaderboards motivated them to complete their homework, and 3 participants mentioned that they were motivated to maintain their rank; these comments about motivation were not present for the middle and low groups. In contrast, from the middle and low group, 2 participants said the leaderboards were initially fun and then became instruments of pressure, 1 participant said the point system should have better reflected the quality of work, not just quantity, and 1 participant said they did not care about the leaderboards; these negative comments were not present for high group. Participants from both groups had differing negative opinions towards not being able to improve their rank. Participant 22 from the high group said that he wished he could have done extra work in order to become the only number one ranked participant, whereas Participant 17 in the low group wrote that it was impossible to improve her overall ranking due to the scoring system, and she could only perform well on the weekly leaderboard.

Table 4.23
Summary of Interview Responses about Leaderboards

Response	Final leaderboard rank		
	high (n = 5)	mid & low (n = 4)	total (n = 9)
+ Score: provided good feedback	2	2	4
+ General motivation: motivated me to complete all homework	3	0	3
+ Rank: motivated me to maintain rank	3	0	3
+ Score: system was motivating	1	1	2
~ Social comparison: pressure to complete homework	2	3	5
~ Weekly schedule was hard when busy	1	1	2
- Negative towards not being able to climb rank	1	2	3
- Leaderboards were fun at first, then became pressure	0	2	2
- Point system could be better	0	1	1
- Apathy towards leaderboard	0	1	1

Note. + = positive comment, ~ = neutral comment, - = negative comment.

4.6 The Language Learning Orientation Scale Results

4.6.1 Reliability of the LLOS

The LLOS was administered to all the participants (N = 46) as a pre-test in Week 1 and as a post-test in Week 14 to examine the effect of leaderboards and quests on L2/FL motivation. For Class 1, 24 participants completed the pre-test and all 26 participants completed the post-test. For Class 2, all 20 participants completed both the pre- and post-tests. Based on Kline's (1999) criteria for evaluating internal consistency in which an alpha greater than .6 is acceptable, Noels et al. (2000) previously demonstrated that the LLOS is a reliable research tool. For this study, the Cronbach's alpha scores were calculated for each subscale using SPSS 24.0. Table 4.24 shows the reliability coefficients. Based on Noels et al. (2000) and the results in Table 4.24, the LLOS was considered a reliable instrument to measure L2/FL motivation for this study.

Table 4.24
Cronbach's α Coefficients for the LLOS Subscales

Scale	# of items	Cronbach's α (pre) *n* = 44	Cronbach's α (post) *n* = 46
Amotivation	3	.18	.45
EM: External regulation	3	.84	.64
EM: Introjected regulation	3	.50	.68
EM: Identified regulation	3	.63	.60
IM: Accomplishment	3	.71	.81
IM: Knowledge	3	.84	.85
IM: Stimulation	3	.77	.80
Extrinsic combined	9	.79	.73
Intrinsic combined	9	.85	.90

Note. EM = extrinsic motivation, IM = intrinsic motivation.

4.6.2 LLOS Paired Samples *t*-Test Results

Paired samples *t*-tests were conducted to compare the pre- and post-test mean scores for each LLOS subscale to determine if L2/FL motivation changed over the duration of the study. Nine subscales were analysed: the seven subscales of LLOS, the combined extrinsic motivation subscale, and the combined intrinsic motivation subscale. For each subscale, three paired samples *t*-tests were conducted: Class 1, Class 2, Class 1&2 combined. A total of 27 paired samples *t*-tests were conducted. Two participants for Class 1 were removed from the analyses as they did not complete the pre-test. Table 4.25 presents the combined descriptive statistics for the 27 paired samples *t*-tests. Table 4.26 presents the combined results of the 27 paired samples *t*-tests. Statistically significant results were observed at the $p < .01$, $p < .05$, and $p < .10$ levels. Figure 4.9 provides a visual representation that compares the changes in subscale mean scores for each of the classes.

Table 4.25
Comparison of LLOS Subscale Mean Scores for all Groups

LLOS subscales	Class	Pre-test		Post-test	
		M	*SD*	*M*	*SD*
Amotivation	1 (*n* = 24)	1.11	.21	1.54**	.77
	2 (*n* = 20)	1.43	.68	1.70	.89
	1&2 (*N* = 44)	1.25	.50	1.61***	.82
EM: External regulation	1	3.19	1.23	3.51	1.31
	2	3.65	1.59	3.55	1.33
	1&2	3.40	1.41	3.53	1.30
EM: Introjected regulation	1	2.06	1.00	2.22	1.18
	2	2.16	.93	2.23	.90
	1&2	2.11	.96	2.22	1.05
EM: Identified regulation	1	5.05	1.43	4.61	1.11
	2	4.98	.88	4.98	1.12
	1&2	5.02	1.20	4.78	1.12
IM: Accomplishment	1	2.97	1.27	3.33	1.46
	2	3.10	1.09	3.20	1.33
	1&2	3.03	1.18	3.27	1.39
IM: Knowledge	1	4.52	1.55	4.54	1.47
	2	4.30	1.45	4.70*	1.65
	1&2	4.42	1.49	4.61	1.54
IM: Stimulation	1	3.62	1.63	3.95	1.27
	2	3.85	1.15	4.28	1.47
	1&2	3.72	1.42	4.10**	1.35
Combined EM subscale	1	3.43	.95	3.44	.91
	2	3.60	.93	3.58	.82
	1&2	3.51	.93	3.51	.87
Combined IM subscale	1	3.70	1.14	3.98	1.26
	2	3.75	1.11	4.06	1.27
	1&2	3.72	1.12	3.99*	1.25

*** = $p < .01$, ** = $p < .05$, * = $p < .10$

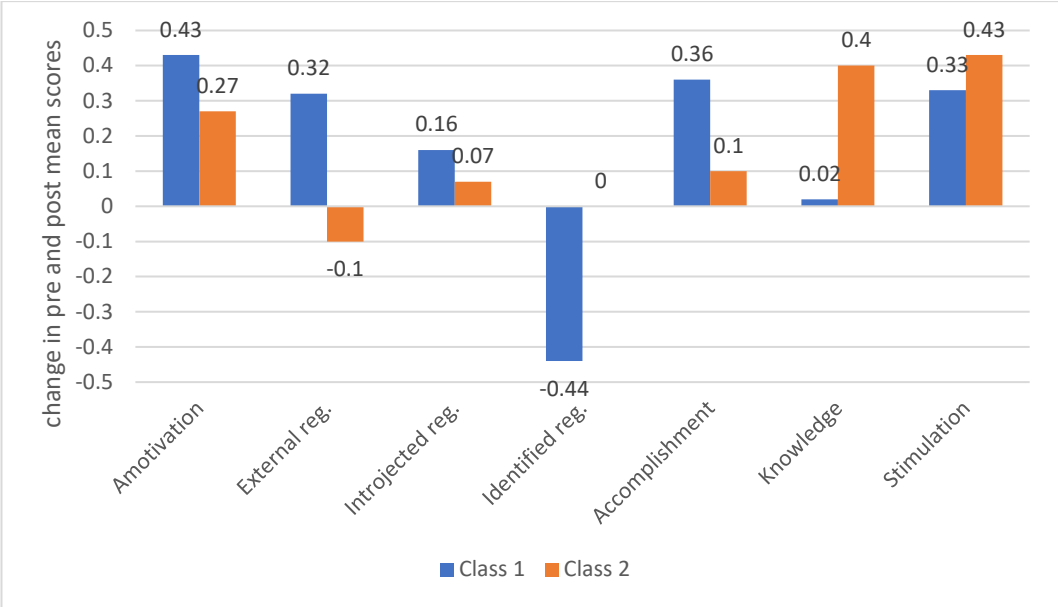


Figure 4.9. Change of mean score for each LLOS subscale for each class.

Table 4.26
LLOS Paired Samples t-test Results for all Groups

		Paired Differences							
					95% Confidence		Interval of difference		
Pair	Class	*M*	*SD*	SEM	Lower	Upper	t	*df*	Sig. (2-tailed)
Pair 1	1	-.43	.81	.16	-.77	-.08	-2.59	23	.016**
Amot_pre-	2	-.26	.68	.15	-.58	.05	-1.73	19	.100
Amot_post	1&2	-.35	.75	.11	-.58	-.12	-3.12	43	.003***
Pair 2	1	-.31	1.02	.20	-.75	.11	-1.52	23	.140
ExtReg_pre-	2	.10	1.07	.23	-.40	.60	.41	19	.681
ExtReg_post	1&2	-.12	1.05	.15	-.44	-.12	-.81	43	.422
Pair 3	1	-.15	1.11	.22	-.62	.31	-.67	23	.509
Introj_pre-	2	-.06	1.03	.23	-.55	.41	-.28	19	.776
Introj_post	1&2	-.11	1.06	.16	-.43	.19	-.70	43	.484
Pair 4	1	.44	1.37	.27	-.13	1.02	1.58	23	.126
Ident_pre-	2	.00	.91	.20	-.42	.42	.00	19	1.000
Ident_post	1&2	.24	1.19	.17	-.12	.21	1.34	43	.185
Pair 5	1	-.36	1.26	.25	-.89	.17	-1.39	23	.176
Accomp_pre-	2	-.10	1.49	.33	-.79	.59	-.299	19	.768
Accomp_post	1&2	-.24	1.36	.20	-.65	.60	-1.17	43	.245
Pair 6	1	-.01	1.51	.30	-.65	.62	-.04	23	.964
Know_pre-	2	-.40	.92	.20	-.83	.03	-1.94	19	.067*
Know_post	1&2	-.18	1.27	.19	-.57	.17	-.98	43	.331
Pair 7	1	-.33	1.25	.25	-.86	.19	-1.30	23	.206
Stim_pre-	2	-.43	1.20	.26	-.99	.13	-1.60	19	.124
Stim-post	1&2	-.37	1.21	.18	-.74	.19	-2.06	43	.045**
Pair 8	1	-.00	.90	.18	-.38	.37	-.05	23	.960
Ext_pre-	2	.01	.83	.18	-.38	.40	.059	19	.953
Ext_post	1&2	.00	.86	.13	-.26	.26	.00	43	1.000
Pair 9	1	-.23	.99	.20	-.65	.18	-1.16	23	.257
Int_pre-	2	-.31	.94	.21	-.75	.13	-1.46	19	.159
Int_post	1&2	-.27	.96	.14	-.56	.02	-1.85	43	.070*

*** = $p < .01$, ** = $p < .05$, * = $p < .10$

The results of the paired samples *t*-tests for the Amotivation subscale show a rise in mean scores for all three groups. For Class 1, there was a significant difference in the scores for Amotivation pre-test ($M = 1.11$, $SD = .21$) and Amotivation post-test ($M = 1.54$, $SD = .77$), $t(23) = -2.59$, $p = .016$. For Class 2, there was no significant difference in the scores for Amotivation pre-test ($M = 1.43$, $SD = .68$) and

Amotivation post-test (*M* = 1.70, *SD* = .89), *t*(19)= -1.73, *p* = .100. For Class 1&2 combined, there was a significant difference in the scores for Amotivation pre-test (*M* = 1.25, *SD* = .50) and Amotivation post-test (*M* = 1.61, *SD* = .82), *t*(43)= -3.12, *p* = .003. These results suggest that the participants' amotivation to study English increased for all groups. The result for Class 1&2 combined was statistically significant at the .01 level, and the result for Class 1 was statistically significant at the $p < .05$ level.

The results of the paired samples *t*-tests for the External regulation subscale show mixed results for the three groups. Mean scores declined for Class 2 but increased for Class 1 and Class 1&2 combined. For Class 1, there was no significant difference in the scores for External regulation pre-test (*M* =3.19, *SD* = 1.23) and External regulation post-test (*M* = 3.51, *SD* = 1.31), *t*(23)= -1.52, *p* = .140. For Class 2, there was no significant difference in the scores for External regulation pre-test (*M* = 3.65, *SD* = 1.59) and External regulation post-test (*M* = 3.55, *SD* = 1.33), *t*(19)= .41, *p* = 681. For Class 1&2 combined, there was no significant difference in the scores for External regulation pre-test (*M* = 3.40, *SD* = 1.41) and External regulation post-test (*M* = 3.53, *SD* = 1.30), *t*(43)= -.81, *p* = .422. None of the results were statistically significant. However, the results suggest that the participants in Class 1 experienced a rise in external regulation, whereas the participants in Class 2 experienced a decline in external regulation.

The results of the paired samples *t*-tests for the Introjected regulation subscale show a rise in mean scores for all three groups. For Class 1, there was no significant difference in the scores for Introjected pre-test (*M* = 2.06, *SD* = 1.00) and Introjected post-test (*M* = 2.22, *SD* = 1.18), *t*(23)= -.67, *p* = .509. For Class 2, there was no significant difference in the scores for Introjected pre-test (*M* = 2.16, *SD* = .93) and Introjected post-test (*M* = 2.23 , *SD* = .90), *t*(19)= -.28, *p* = .776 . For Class 1&2 combined, there was no significant difference in the scores for Introjected pre-test (*M* = 2.11, *SD* = .96) and Introjected post-test (*M* = 2.22, *SD* = 1.05), *t*(43)= -.70, *p* = .484. These results suggest that there was a slight rise in introjected regulation for all groups; however, none of the results were significant.

The results of the paired samples *t*-tests for the Identified regulation subscale show

mixed results for the three groups. The mean scores for Class 1 declined, stayed the same for Class 2, and declined for Class 1&2 combined. For Class 1, there was no significant difference in the scores for Identified pre-test (M = 5.05, SD = 1.43) and Identified post-test (M = 4.61, SD = 1.11), $t(23) = 1.58$, p = .126. For Class 2, there was no significant difference in the scores for Identified pre-test (M = 4.98, SD = .88) and Identified post-test (M = 4.98, SD = 1.12), $t(19) = 0.00$, p = 1.000. For Class 1&2 combined, there was no significant difference in the scores for Identified pre-test (M = 5.02, SD = 1.20) and Identified post-test (M = 4.78, SD = 1.12), $t(43) = 1.34$, p = .185. These results suggest that identified regulation declined for Class 1 and remained the same for Class 2; however, none of the results were statistically significant.

The results of the paired samples *t*-tests for the Accomplishment subscale show a rise in mean scores for all three groups. For Class 1, there was no significant difference in the scores for Accomplishment pre-test (M = 2.97, SD = 1.27) and Accomplish post-test (M = 3.33, SD = 1.46), $t(23) = -1.39$, p = .176. For Class 2, there was no significant difference in the scores for Accomplishment pre-test (M = 3.10, SD = 1.09) and Accomplish post-test (M = 3.20, SD = 1.33), $t(19) = -.299$, p = .768. For Class 1&2 combined, there was no significant difference in the scores for Accomplishment pre-test (M = 3.03, SD = 1.18) and Accomplish post-test (M = 3.27, SD = 1.39), $t(43) = -1.17$, p = .245. These results suggest that all the participants' intrinsic motivation to perform well using English increased; however, none of the results were statistically significant.

The results of the paired samples *t*-tests for the Knowledge subscale show a rise in mean scores for all three groups. For Class 1, there was no significant difference in the scores for Knowledge pre-test (M = 4.52, SD = 1.55) and Knowledge post-test (M = 4.54, SD = 1.47), $t(23) = -.04$, p = .964. For Class 2, there was a significant difference in the scores for Knowledge pre-test (M = 4.30, SD = 1.45) and Knowledge post-test (M = 4.70, SD = 1.65), $t(19) = -1.94$, p = .067. For Class 1&2 combined, there was no significant difference in the scores for Knowledge pre-test (M = 4.42, SD = 1.49) and Knowledge post-test (M = 4.61, SD = 1.54), $t(43) = -.98$, p = .331. These results suggest that there was an increase in all the participants' intrinsic motivation to improve their English skills because of the resulting positive

feelings. However, statistical significance ($p < .10$) was only registered for Class 2.

The results of the paired samples *t*-tests for the Stimulation subscale show a rise in mean scores for all three groups. For Class 1, there was no significant difference in the scores for Stimulation pre-test ($M = 3.62$, $SD = 1.63$) and Stimulation post-test ($M = 3.95$, $SD = 1.27$), $t(23) = -1.30$, $p = .206$. For Class 2, there was no significant difference in the scores for Stimulation pre-test ($M = 3.85$, $SD = 1.15$) and Stimulation post-test ($M = 4.28$, $SD = 1.47$), $t(19) = -1.60$, $p = .124$. For Class 1&2 combined, there was a significant difference in the scores for Stimulation pre-test ($M = 3.72$, $SD = 1.42$) and Stimulation post-test ($M = 4.10$, $SD = 1.35$), $t(43) = -2.06$, $p = .045$. These results suggest that there was an increase in all the participants' intrinsic motivation to engage in the process of learning English because of the associated positive feelings; however, statistical significance ($p < .05$) was only registered for Class 1&2 combined.

The results of the paired samples *t*-tests for the combined extrinsic motivation subscale show similar mean scores for all three groups. For Class 1, there was no significant difference in the scores for Extrinsic pre-test ($M = 3.43$, $SD = .95$) and Extrinsic post-test ($M = 3.44$, $SD = .91$), $t(23) = -.05$, $p = .960$. For Class 2, there was no significant difference in the scores for Extrinsic pre-test ($M = 3.60$, $SD = .93$) and Extrinsic post-test ($M = 3.58$, $SD = .82$), $t(19) = .05$, $p = .953$. For Class 1&2 combined, there was no significant difference in the scores for Extrinsic pre-test ($M = 3.51$, $SD = .93$) and Extrinsic post-test ($M = 3.51$, $SD = .87$), $t(43) = .00$, $p = 1.000$. These results suggest that all the participants' extrinsic motivation towards studying English remained the same for all three groups.

The results of the paired samples *t*-tests for the combined intrinsic motivation subscale show a rise in mean scores for all three groups. For Class 1, there was no significant difference in the scores for Intrinsic pre-test ($M = 3.70$, $SD = 1.14$) and Intrinsic post-test ($M = 3.98$, $SD = 1.26$), $t(23) = -1.16$, $p = .257$. For Class 2, there was no significant difference in the scores for Intrinsic pre-test ($M = 3.75$, $SD = 1.11$) and Intrinsic post-test ($M = 4.06$, $SD = 1.27$), $t(19) = -1.46$, $p = .159$. For Class 1&2 combined, there was a significant difference in the scores for Intrinsic pre-test ($M = 3.72$, $SD = 1.12$) and Intrinsic post-test ($M = 3.99$, $SD = 1.25$), $t(43) = -1.85$, $p = .070$.

These results suggest that intrinsic motivation increased for all groups. However, statistical significance ($p < .10$) was only registered for Class 1&2 combined.

4.7 Summary

This chapter has presented the results of the various data analyses that were conducted on the data collected from the five data collection instruments. This section summaries the results of data analyses. Section 4.2 has presented the results of the independent samples *t*-tests were conducted on the performance-related data to determine if there were any differences in performance between Class 1 and Class 2 for the six measurements of performance. The results showed no statistically significant differences in performance between Class 1 and Class 2. However, the descriptive statistics showed that the mean scores for the English Central activities and the number of quests completed were slightly higher for Class 2, and the mean scores for the two MReader performance measures were noticeably lower for Class 1.

Section 4.3 has presented the results of the data analyses that were conducted on the leaderboard questionnaire data. The results showed that the high leaderboard ranked participants more frequently selected positive emotions such as determined and enjoyment, whereas the low ranked participants more frequently selected the negative emotions such as anxiety and shame. The results suggested that the point and rank system was beneficial for the high ranked participants but detrimental for the low ranked participants.

Section 4.4 has presented the results of the data analyses that were conducted on the quest diaries. The results showed overwhelming positive opinions and perceptions towards the quests and the QBL experience. The participants enjoyed the experience and they felt it was a good approach to learning EFL. The participants appreciated being able to choose which quest to work on. They often chose a quest based on how easily or quickly they thought they could finish it, how enjoyable it looked, or because they wanted to improve or use their English skill through the quest. The results showed that the participants felt positively about the collaboration aspect of

questing because it made the activities more enjoyable and they could spend meaningful time with people. However, some of the participants avoided collaborating because they wanted to work by themselves or because they did not want to work with other people.

Section 4.5 has presented the results of the content analysis that was conducted on the semi-structured interviews. The results of the analysis showed a range of opinions towards the quests and the leaderboards. Regarding the quests, the most common comments were that having a choice of quest to do was good, the quests were initially difficult but became easier, and the social aspect of the quests was enjoyable. Regarding the leaderboards, the most common comments were that the continued social comparison pressured the participants to complete their homework, the score provided good feedback, the high ranked participants were motivated to maintain their rank, and the rank and point system need to be adjusted.

Section 4.6 has presented the results of the paired samples *t*-tests that were conducted on the pre- and post-test LLOS subscales to determine if the participants' FL motivation changed over the duration of the study. The results showed that, for the extrinsic motivation subscales, external regulation increased for Class 1 but decreased for Class 2, and identified regulation decreased for Class 1 but stayed the same for Class 2. For the intrinsic motivation subscales, accomplishment increased for both classes, knowledge increased significantly for Class 2, and stimulation increased for both classes. When the three extrinsic motivation subscales were combined, the pre- and post-test mean scores remained the same for both classes. However, when the three intrinsic motivation subscales were combined, the mean scores increased for both classes with the result being statistically significant for Class 1&2.

Chapter 5. Discussion

5.1 Overview

This chapter discusses the results reported in Chapter 4 in relation to the research questions. The results of the individual data collection instruments are first discussed in isolation, and then combined with the results of the other instruments to triangulate findings for the research questions. The findings are then compared to the existing literature. The chapter concludes with a summary of the main findings.

5.2 The Effect of Leaderboards on Performance

This section discusses the effect the leaderboard had on the participants' performance. Section 5.2.1 discusses and compares the performance-related data between Class 1 and Class 2. Section 5.2.2 discusses the results of the leaderboard questionnaire with a focus on the participants' emotions and attitudes towards the leaderboards. The final sections combine the results of the performance-related data, the leaderboard questionnaires, and the semi-structures interviews to discuss how the participants' performance was affected by the point and rank components of the leaderboard.

5.2.1 The Performance-related Data Findings

The performance-related data suggest that the leaderboard affected performance. Even though the results of the independent samples *t*-tests that compared the final mean scores of Class 1 and Class 2 for each measure of performance reported no statistically significant differences, the descriptive statistics provide interesting insight. The mean scores show that Class 2 basically achieved all five of their final target point totals, whereas, Class 1 only achieved four of their final target point totals. The performance-related data show that Class 2 recorded higher mean scores for all measures of performance; slightly higher for four of the performance measures (i.e., EC_Videos_watched, EC_videos_spoken, EC_goals_completed, and Quests_completed), and noticeably higher for two of the measures (i.e.,

MR_quizzes_passed and MR_words_read).

To further analyse the data, the participants' data from each class were categorised based on their performance as either low, middle, or high (explained in Section 4.2). The data show that the high and middle ranked participants in both classes basically achieved their weekly goals; however, the low ranked participants did not. The class comparison of the low group data (see Table 4.3) show very similar results; both low groups were occupied by about 21% of each class, and both low groups recorded similar performance scores for all measures except MR_words read. The average amount of MR_words read for Class 1 was much lower than Class 2. The class comparison of the middle group data show that 69% of the participants in Class 1 are in the middle group and 50% of the participants in Class 2 are in the middle group. The class comparison of the high group data show that 30% of the participants in Class 2 are in the high group whereas only 8% of the participants in Class 1 are in the high group. The middle and high group data could suggest that most of the participants in Class 1 stopped working when they had achieved their weekly point goals, however, many of the participants in Class 2 continued working even when they had achieved their weekly point goals.

The only clear finding resulting from the performance-related data is that the leaderboard impacted the performance of the participants in Class 1 by limiting performance. The finding cannot support the literature (e.g., Aldemir et al., 2018; Borys & Laskowsky, 2013; Domínguez et al., 2013; Halan, Rossen, Cendan, & Lok, 2010; Landers et al., 2015; Mekler et al., 2013; Tan & Hew, 2016) that found that leaderboards positively impact learner behaviour. However, the finding cannot refute the literature because it is possible that, if leaderboards were not present in Class 1, the amount of work completed could have been much less. The finding partially supports Bursztyn and Jensen (2015) who found that a leaderboard led to a decrease in performance; the finding only provides partial support because even though the outcomes were similar, the reasons for the outcomes appear different. Bursztyn and Jensen (2015) attributed the decrease in performance to students wanting to avoid appearing on a top 3 leaderboard, whereas the results of the performance-related data suggest that the limited performance was due to the participants ceasing work once they had achieved the maximum point reward offered by the leaderboard.

5.2.2 The Leaderboard Questionnaire Findings

The leaderboard questionnaire collected data about the participants' emotions, attitudes, and opinions towards the leaderboards at Week 4, 8, and 12. In the emotions section of the leaderboard questionnaire, the participants selected the emotions they felt when they looked at the weekly and overall leaderboards, then explained why they made that decision. The results of the leaderboard questionnaire show no noticeable differences between how the participants responded to the weekly and overall leaderboards. This went against the researcher's expectations as he thought the participants who were not performing well on the overall leaderboard would show fewer positive emotions towards the overall leaderboard as the chance of gaining a high ranking declined as the semester progressed. To simplify the discussion about the self-reported emotions, this section uses the average frequency of the weekly and overall emotion data. The data show that about 40% of the participants' responses selected multiple emotions to reflect their feeling towards the leaderboards. This supports the literature (e.g., Izard, 1991; Zeldin, 1995) that warns that understanding emotions is difficult because they do not occur in isolation. Even though the performance-related data show better results for Class 2, the results of the leaderboard questionnaire suggest that the participants in Class 1 responded well to the leaderboard.

The results from the leaderboard questionnaire suggest that most of the participants in Class 1 responded positively to the leaderboard. The results of the emotion section show that, when the participants looked at the leaderboard, they reported the positive emotions enjoyment, hope, pride, and determined noticeably more often than the negative emotions anxiety, shame, hopelessness, and envy. When the participants chose a positive emotion, they were more likely to achieve a full performance score in the following week. The participants explained that they chose the positive emotions because they were motivated by the ranking system, they were responding positively to an aspect of the point system, and they felt a sense of accomplishment through the leaderboards. The results of the attitude section show that, when the participants looked at the leaderboard, they had positive attitudes towards maintaining and gaining rank, improving their class grade and English ability, and the leaderboard being fun and something they care about. The results of the opinion

section of the leaderboard questionnaire also show a high frequency of positive opinions towards the leaderboard being good for motivation, and a low frequency of negative opinions. The combination of these results shows a connection between positive emotions, positive performance, and provides reasons why positive emotions lead to positive performance.

The most frequently self-reported emotion that the participants felt when they looked at the leaderboard was determined; determined was chosen noticeably more often than any of the other emotions. It was chosen twice as often as the second most frequently selected emotion, enjoyment. On average, determined was reported by about 16 participants (62%) for each of the three surveys. Ranked from most frequent to least frequent, the participants explained that the reasons for choosing determined were: (1) they were motivated by the ranking system, (2) they were responding positively to an aspect of the point system, (3) they felt a sense of accomplishment through the leaderboards. The results of the attitude section of the leaderboard questionnaire and the semi-structured interviews confirm the positive attitudes and opinions towards the rank and score system. When the participants chose the determined emotion, 85% of the time they went on to get a full score in the following week. The positive results surrounding the determined emotion not only support the literature (i.e., Aldemir et al., 2018; Borys & Laskowsky, 2013; Domínguez et al., 2013; Halan, Rossen, Cendan, & Lok, 2010; Landers et al., 2015; Mekler et al., 2013; Tan & Hew, 2016) that found that leaderboards positively impact performance, but further show that it is the rank and score system that affect performance.

The second most frequently self-reported emotion was enjoyment; enjoyment was chosen over twice as often as the third most frequently selected emotion, anxiety. On average, enjoyment was reported by about nine participants (33%) for each of the three surveys. Ranked from most frequent to least frequent, the participants explained that the reasons for choosing enjoyment were: (1) they were motivated by the ranking system, (2) they felt a sense of accomplishment through the leaderboards, (3) they were responding positively to an aspect of the point system. 94% of the time that enjoyment was selected, the participants went on to get a full score in the follow week. The attitude results show that the participants slightly agreed that the

leaderboard was fun. The frequency that enjoyment was selected lightly supports the gamification literature (e.g., De Schutter & Abeele, 2014; Dubravac, 2012; Domínguez et al., 2013; Jia et al., 2017; Wells & Skowronski, 2012) that asserts students enjoy leaderboards.

The fourth most frequently self-reported emotion was hope; on average, hope was reported by about three participants (10%) for each of the three surveys. Ranked from most frequent to least frequent, the participants explained that the reasons for choosing hope were: (1) they were responding positively to an aspect of the point system; (2) they felt a sense of accomplishment; and (3) they were motivated by the ranking system. 92% of the time that hope was selected, the participants went on to get a full score in the following week. The fifth most frequently self-reported emotion was pride; on average, pride was reported by about three participants (10%) for each of the three surveys. Ranked from most frequent to least frequent, the participants explained that the reasons for choosing pride were: (1) they felt a sense of accomplishment; (2) they were motivated by the ranking system; and (3) they were responding positively to an aspect of the point system. 100% of the time that pride was selected, the participants went on to get a full score in the follow week. The low frequency that hope and pride were selected in comparison to determined and enjoyment suggest that they are of less importance in this analysis of the effect leaderboards have on performance. However, they do support the connection between choosing positive emotions and positive performance.

Data collected from the leaderboard questionnaire suggest that the leaderboard had a negative impact on a small number of the participants. The third most frequently self-reported emotion was anxiety; on average, anxiety was reported by about four participants (14%) for each of the three surveys. Ranked from most frequent to least frequent, the participants explained that the reasons for choosing anxiety were: (1) they felt negatively about their rank; (2) they felt negatively about the social comparison; and (3) they felt negatively about their score. 79% of the time that anxiety was selected, the participants went on to get a full score in the follow week. These results are similar to Aldemir et al. (2018) who found that deadlines in a gamified course lead to anxiety for some students.

The data collected about the other negative emotions show that on average, two participants (7%) chose shame, and one participant (3%) chose hopelessness for each of the three surveys. The participants explained that they chose shame and hopelessness due to negative feelings towards their rank, score, and the social comparison. 0% of the time that shame was selected, and 33% of the time that hopelessness was selected, the participants went on to get a full score in the follow week. The results surrounding the selection of the negative emotions anxiety, shame, and hopelessness support the literature (e.g., Barata, Gama, Jorge, & Gonçalves, 2013; Charles et al., 2011; Domínguez et al., 2013; Nicholson, 2013) that found some students have negative attitudes towards the competitive and the comparative nature of leaderboards, and further show that a negative emotional reaction is less likely to lead to a positive performance outcome compared to a positive emotional reaction.

5.2.3 The Point System Affects Performance

The data suggest that the point system had an effect on the participants' performance. The leaderboard questionnaire data show that, when the participants in Class 1 explained why they chose an emotion, the third most frequent reason was due to their positive opinion of the score system. The participants explained that the point system motivated them by providing goals, and feedback about their performance. The semi-structured interview data also show that over half of the participants mentioned that the point scoring system motivated them to work harder and provided them with good feedback. The performance-related data (presented in Table 4.3) show that most of the participants in Class 1, except for the low performing participants, did just about enough work to achieve their final target point totals but not much more. The combination of the leaderboard questionnaire results and the performance-related data suggest that the participants were motivated to achieve their weekly point goals. This finding supports the literature (i.e., Aldemir et al., 2018; Domínguez et al., 2013; Iosup & Epema, 2014; Tan & Hew, 2016) that found that gamification components, such as point systems, have a positive effect on performance by making the participants more cognisant of the type of behaviour that will be rewarded, and by providing clear goals and the ability to self-assess

performance. However, the finding comes with two caveats discussed in the paragraphs below.

The first caveat is that the point system limits performance. The performance-related data (presented in Table 4.3) show the presence of a large group of high performing participants in Class 2 who completed noticeably more work than the point system rewarded. In contrast the size of the high performing group in Class 1 was much smaller.In the semi-structured interviews, three participants expressed negativity about not being able to climb the leaderboard due to the limitations of the point system; the participants would have completed more work if the point system did not stipulate a weekly maximum. The results suggest that maybe the high ranked participants in Class 1 would have done more work if they were not constantly exposed to the point system.

The learning theory of behaviourism (explained in Section 2.4.1) explains how leaderboards use extrinsic rewards such as points to encourage participants to behave in a desired way. However, the behaviour must be continually rewarded, or else compliance will cease. These results suggest that most of the participants in Class 1 were motivated to achieve their final target point total but they stopped completing the homework activities once they had secured the maximum amount of points they could be rewarded with. The point system represented the minimum amount of work the teacher wanted the participants to complete. Therefore, even though the point system successfully influenced performance up to a certain point, it also negatively impacted performance by creating a reward ceiling. The finding that the point system limits behaviour supports the gamification literature (e.g., Borys & Laskowsky, 2013; Kapp, 2012; Koivisto & Hamari, 2014; Krause, Mogalle, Pohl, & Williams, 2015; Mollick & Rothbard, 2014; Sheldon, 2011) that asserts that extrinsic goals create a performance ceiling and, once the goal has been achieved, the participant is unlikely to continue doing the behaviour if it is not rewarded.

The second caveat is that point system could negatively impact the quality of the learning outcomes. The data collected from the leaderboard questionnaires and the semi-structured interviews show a few instances of the participants commenting that the point system should have better reflected the quality of work, not just quantity. If

the participants were only focused on doing the learning activities to get points, rather than trying to improve their English ability, it would not lead to an ideal learning outcome. Even though it was only mentioned by a few participants, it does not mean it was not true for other participants. The comments provide light support for the literature (e.g., Domínguez et al., 2013; Halan, Rossen, Cendan, & Lok, 2010; Tan & Hew, 2016; Werbach & Hunter, 2012) that claims that, if a point system prioritises quantity over quality, it may negatively affect the quality of the learning outcomes. The comments also support the warnings (e.g., Aldemir et al., 2018; Hanus & Fox, 2015) that that extrinsic rewards such as points need to reward the desired behaviour, not just behaviour.

5.2.4 Leaderboard Rank Affects Performance

The literature (e.g., Burguillo, 2010; Hakulinen et al., 2013; Orosz et al., 2013; Lam et al., 2004; Reeve & Deci, 1996) explains that competition can increase performance, motivation, and enjoyment of tasks, but can also lead to a range of negative outcomes such as disengagement or an unwanted feeling of pressure. The gamification literature (e.g., Barata et al., 2017; Cheong et al., 2013; Garland, 2015; Hamari et al., 2014; Jia et al., 2017; Philpott, 2015a) warns that the use of a leaderboard in a gamified course will have a a psychological effect on students, and different students will respond differently to it. The results of the leaderboard questionnaires and semi-structured interviews show that leaderboard rank was an influential component of the leaderboard. The results also suggest that leaderboards affect participants differently based on their leaderboard rank; high ranked participants respond positively, low ranked participants respond negatively.

Leaderboards positively affect the emotions, attitudes, and performance of high ranked participants. The leaderboard questionnaire data show that on average, determined was selected by the participants with a high rank 77% of the time, middle rank 80% of the time, and low rank 37% of the time. Enjoyment was selected by the participants with a high rank 55% of the time, middle rank 38% of the time, and low rank 21% of the time. Over the three surveys, the number of participants who chose the pride emotion continued to rise for the high ranked group but remained at 0 for

the low ranked group. When the participants chose a positive emotion, they were more likely to achieve a full performance score in the following week. The high-ranked participants also showed noticeably more positive attitudes towards the leaderboards compared to the low-ranked participants. For the Activities_neg attitude, the final mean score was noticeably lower for the high ranked group compared to the low ranked group. The results of the ANOVA found statistically significant differences between the high ranked group and the low ranked group for the Maintain_rank and LB_is_fun attitudes. They suggest that, compared to the low ranked group, the high ranked participants were motivated to work hard to maintain their rank, they felt that the leaderboard was fun, and they felt more positively about the learning activities.

The difference in attitudes based on leaderboard rank was also noted in the semi-structured interviews. Only the high ranked participants said that they were motivated by the leaderboard in general, and the rank aspect of the leaderboard. The combination of the results discussed in this section clearly suggest that a leaderboard is more likely to have a positive effect on high ranked participants. The finding supports the literature (e.g., Aldemir et al., 2018; Buckley, Doyle, & Doyle, 2017; Çakıroglu et al., 2017; Heeter, Lee, Medler, & Magerko, 2011; Tan & Hew 2016; Tran & Zeckhauser 2012; Werbach & Hunter, 2012) that explains that a competitive social leaderboard positively affects the performance of higher performing students.

Leaderboards negatively affect the emotions, attitudes, and performance of low ranked participants. The leaderboard questionnaire data show that the lower ranked participants reported more negative emotions and attitudes towards the leaderboard. The data show that the third most frequently self-reported emotion anxiety, on average, was reported by 33% of the low-ranked participants, 8% of the middle-ranked participants, and 11% of the high-ranked participants. The participants explained that the feeling of anxiety stemmed from negative feelings towards their rank, their score, or the forced social comparison. The negative emotions shame and hopelessness were basically only chosen by low ranked participants; on average over the three surveys, shame was chosen 25% of the time by the low ranked participants, and hopelessness was chosen 16% of the time. The participants explained that they chose shame and hopelessness because they had lost confidence and that their rank or

score was too bad. When the participants chose a negative emotion, they were more likely to not achieve a full performance score in the following week.

The leaderboard questionnaire data show noticeably more negative attitudes towards the leaderboards for the low-ranked participants compared to the high ranked participants. They show noticeably different mean scores for the Activities_neg attitude, and statistically significant differences for the Maintain_rank and LB_is_fun attitudes. They suggest that compared to the high ranked participants, the participants in the low group were not motivated to maintain their rank, they did not consider the leaderboard fun, and they felt more negatively towards the learning activities. For the attitude that represents not caring about the leaderboard, there is a noticeable difference in mean score between the high ranked group and the low ranked group; the mean score for the low group suggests that they slightly did not care about the leaderboard. In the semi-structured interviews, two participants (50%) who were not in the high ranked group said the leaderboards were initially fun but became forces of negative pressure.

The finding that the rank aspect of a leaderboard negatively affects some of the participants supports the literature (e.g., Barata, Gama, Jorge, & Gonçalves, 2013; Charles et al., 2011; Domínguez et al., 2013; Nicholson, 2013) that explains that the competitive and comparative aspects of a leaderboard in an education setting lead to negative feelings, and further shows that the negative feelings affect performance. The finding that low performing participants are more likely to respond negatively to a leaderboard that shows their rank supports the literature (e.g., Aldemir et al., 2018; Buckley, Doyle, & Doyle, 2017; Çakıroglu et al., 2017; Tan & Hew 2016; Werbach & Hunter, 2012) that warns leaderboard rank negatively affects the performance of lower performing students.

The performance-related data and the leaderboard questionnaire data suggest that for the low ranked group, the leaderboard initially had a positive impact, but over time, engagement declined. The performance-related data (presented in Table 4.4) show that performance declined for the low group; for the first three-week period, they received 92% of total points, for the second three-week period, they received 69% of total points, for the third three-week period, they received 78% of total points, and

for the final three week period, they received 67% of total points. In contrast, the high ranked participants received 100% of total points every week. The leaderboard questionnaire data show that the amount of low ranked participants who self-reported the determined emotion declined significantly, especially between Survey 1 and Survey 2 in which the percentage dropped from 75% to 12% of the group. For the low ranked group, the decline in choosing determined corresponded with a rise in choosing anxiety, especially between Survey 1 and Survey 2 in which it increased from 12% to 62%. The results of the Friedman rank sum tests show a statistically significant change for the Don't_care attitude between Survey 1 and Survey 2, and Survey 3. The high final mean score for the low ranked group for the Don't_care attitude suggests that the statistically significant change is due to the low ranked group. The finding suggests that the low ranked participants cared less about the leaderboard at the time of Survey 3, than at the time of Survey 1 and Survey 2.

For the low ranked group, the declines in performance and the selection of the determined emotion, alongside the rises in choosing anxiety and the Don't_care attitude support Fotaris, Mastoras, Leinfellner, and Rosunally's (2016) finding that student engagement declines as leaderboard ranking declines. The finding partially supports the literature (e.g., Borys & Laskowsky, 2013; Buckley, Doyle, & Doyle, 2017; Garland, 2015; Krause, Mogalle, Pohl, & Williams, 2015; Koivisto & Hamari, 2014; Mollick & Rothbard, 2014; Werbach & Hunter, 2012) that warns that gamification techniques, such as leaderboards, are good for short-term performance boosts but will not be able to sustain engagement over an extended period-of-time once the novelty wears off; the finding only partially supports as the decline in performance was only registered for the low ranked participants. Behaviourism (e.g., Chen, 2003) suggests that the reason why the low ranked participants' performance declined and the high ranked participants' performance remained constant was because the high ranked participants continued to be rewarded with high rank, whereas the incentive of high rank was taken away from the low ranked participants as soon as their performance declined. In fairness to gamification, the proponents of gamification (e.g., Chou, 2015; Werbach & Hunter, 2012) clearly explain that gamification implementations that rely on extrinsic rewards will not sustain engagement; they suggest that extrinsic rewards provide direction, and engagement

should be fostered by through incorporating deeper psychological theory, such as SDT, into the design of the implementation.

5.3 Opinions and Perceptions of Quest-based Learning: The Quest Diary Findings

This section discusses the participants' opinions and perceptions of the QBL experience. The discussion centres around the results of the content analysis that was conducted on the quest diaries (presented in Section 4.4); however, the results of the semi-structured interviews support aspects of the discussion when necessary. In general, the results of the content analysis show that the participants had overwhelmingly positive opinions and perceptions of QBL. This section discusses three major themes that emerged from the data: (1) how QBL supports FL acquisition; (2) how QBL supports intrinsic motivation; and (3) how QBL hinders FL motivation and acquisition.

5.3.1 How QBL Supports Foreign Language Acquisition

The results of the quest diary data analysis suggest that the participants perceived QBL as a good approach for learning EFL. For example, when the participants assigned a score to each of the four main learning activities used in the study, the results show that the participants agreed that QBL is good for EFL pedagogy; overall, it was ranked slightly higher than English Central and extensive reading, and noticeably higher than the textbook. Also, when the participants reflected on the QBL experience, the most common perception, with 17 participants (40%) mentioning it, was that QBL is good for EFL pedagogy. This section discusses how further analysis of the quest diary data from an SLA perspective led to three main reasons why QBL was beneficial for FL acquisition: (1) provided many opportunities to use or improve English skills; (2) learned new ways to study EFL; and (3) facilitated meaningful situations to use English.

The first reason why the participants perceived QBL to be good for FL acquisition was because it provided them with many opportunities to use or improve their English skills. In response to the individual quests, the results show that 92% of the participants mention at least once that they could use or improve their English skills during the quest; it was the second most frequent reflection about the individual quests. The reflections of the individual quests provide deeper insight into what English skills the participants said they were using and improving at least once; 60% of the participants mention listening skills, 55% of the participants mention speaking skills, 45% of the participants mention writing skills, 30% mention presentations skills, and 20% mention pronunciation and vocabulary skills. In the semi-structured interviews, three participants also specifically mentioned that QBL was a good way to study English.

The second reason why the participants perceived QBL to be good for FL acquisition was because they learned many new ways to study EFL. In response to the individual quests, the results show that 70% of the participants mention at least once that they learned a new way to study English. The positive perception was also confirmed in the semi-structured interviews; for example, Participant 22 explained that

> "The quests were a good opportunity to learn new tools and websites for learning English. For example, I could learn about websites that had various videos and podcasts for learning English that I could access on my phone. The quests provided me good opportunities to practice my speaking and writing skills. Normally practicing these skills is troublesome for me." (Participant 22)

The third reason why the participants perceived QBL to be good for FL acquisition was because it allowed them to use English in meaningful situations to learn. The analysis of the participants' perceptions of the QBL experience show that the fifth most common perception was that it provided meaningful opportunities; for example,

- Participant 27 wrote: "I learned a lot through the quests. I learned a lot because I learned the content of many subjects, and I learned the English about the subjects."

- Participant 34 wrote: “In the quests, I had to express myself in English. So, I had to search for the words and idioms I required. It was good study.”

The participants also wrote about the non-language related content they learned: 55% of the participants said that they learned I.T. skills, and 15% said that they learned general knowledge. However, the participants more frequently talked about the international knowledge they learned or experienced. 67% of the participants said at least once they learned international knowledge; more specifically, 60% of the participants said they learned about foreign countries, 22% said they learned about different ways of thinking, and 22% said they learned about foreign culture. For example,

- Participant 23 wrote: “It was interesting that I could touch international culture in English through Podcasts and the ELLLO website.”
- Participant 36 wrote: “By working on quests related to international culture, I was able to feeling like I was on study abroad, even though I was still in Japan.”

The general finding that the participants perceived QBL as beneficial for FL acquisition, alongside the three reasons why, provides evidence that the quests were successfully designed for EFL pedagogy. The participants’ comments about the quests allowing them to improve or practice their different English skills, and the comments about learning new ways to study EFL show examples of the quests providing opportunities for comprehensible input (Krashen, 2009) and comprehensible output (Swain, 1985) to occur. Specifically, Participant 34’s comment about having to search for the correct words and idioms to use shows an example of Swain’s Noticing hypothesis in which L2 acquisition occurs when learners realise what they do not know and then they learn what they need to know in order to communicate their opinion.

The participants’ comments about using English in meaningful situations provide examples that the quests provided opportunities for real-life authentic language use which is encouraged in the TBLT literature (e.g., Ellis, 2003; Nunan, 1989; Willis &

Willis, 2007) as a foundation for L2/FL acquisition to occur. Participant 23 and 36s' comments about QBL exemplify how the quests provided meaningful EFL educational experiences in which the participants could learn about the world around them using English. The participants' comments about the international knowledge they learned suggest that international posture (Yashima, 2002; 2009) was succesfully targeted, which is beneficial for fostering FL motivation (Aubrey & Nowlan, 2013; Kim & Lee, 2013; Xie, 2014; Yashima, 2009) in the Japanese EFL context. The finding that the participants had positive perceptions towards QBL as a pedagogical approach not only supports the previous studies (e.g., Çakıroglu et al., 2017; Lambert, et al., 2015) that found students have positive perceptions towards QBL, but further provides three reasons why participants have positive perceptions towards QBL as an approach for EFL pedagogy.

5.3.2 How QBL Supports Intrinsic Motivation

The results of the quest diary data analysis suggest that most of the participants perceived QBL as an intrinsically motivating approach to learning EFL. For example, in response to the individual quests, the results show that the fourth most frequently coded category of response reflects a feeling of motivation; 90% of the participants were coded into the category at least once. The sub categories of the motivation category show what a quest motivated a participant to do at least once; 23 participants (57%) said improve their English skills, 20 participants (50%) said do positive things for their own lives, 12 participants (30%) said do positive things for other people, 11 participants (27%) said get international exposure, and seven participants (17%) said improve their presentation skills. Further analysis of the quest diary data from an SDT and positive psychology perspective provide insight into what aspects of the QBL experience were intrinsically motivating. This section presents five reasons why the QBL experience was an intrinsically motivating pedagogical approach: (1) enjoyable quests; (2) meaningful quests; (3) suitably challenging quests; (4) quest choice supported autonomy; and (5) meaningful interpersonal experiences.

Most of the participants perceived the QBL experience to be enjoyable. The

participants were not specifically asked if they enjoyed the QBL experience or not; however, the results show various instances of the participants conveying that they enjoyed the experience. First, in response to the individual quests, the most frequently coded category represents the various general positive feelings the participants had towards a quest; 24% of all the 909 comments were coded to the category, with 95% of the participants coded to the category at least once. Subsumed under the general positive feelings category are two high frequency subcategories. The subcategory that has the highest number of individual participants coded to it represents the opinion that a quest was enjoyable; 82% of the participants said it at least once. For example, in her reflection about Quest 10, Participant 1 said, "I watched video news for this quest and I found there were many other interesting videos that were not related to the quest, so I watched them too. It was interesting to see news about Japan from a non-Japanese perspective."

Second, the semi-structured interview data show that many of the participants mentioned that the social and creative aspects of QBL were enjoyable. For example, in her semi-structured interview, Participant 1 explained that

> "In the beginning, I only did the writing quests because they were quick to finish, but I gradually started working on videos and discussion quests. I was reluctant to record a video at first, but I tried it and it was really fun and I was able to enjoy studying." (Participant 1)

Finally, the results of the quest diaries show that when the participants shared their reflective opinions about the class in general, the tied most frequent opinion, with 21 participants (49%) mentioning it, was that the class was enjoyable, interesting, or fun. Even though the results about the class in general do not specifically target QBL, the quests were a major component of the course that could not be ignored.

Both SDT and the positive psychology literature (e.g., Seligman, 2008) explain that positive emotions like enjoyment lead to, and are signs of, intrinsic motivation towards the activity domain. SDT suggests that enjoyment occurs when someone is engaged in an activity that fosters competence, autonomy, and relatedness. Lambert et al. (2015) found that students enjoy QBL because it was suitably challenging. The finding from this study that the participants perceived QBL as an enjoyable

experience supports Lambert et al. (2015) and further shows how the autonomous-nature of QBL leads to enjoyment if the participants can explore content they are interested in, without being forced.

Most of the participants perceived the QBL experience to be meaningful. The positive psychology literature (e.g., Seligman, 2008) explains that intrinsic motivation can foster through activities that are personally meaningful. The participants were not specifically asked if the quests were meaningful; however, the results show various instances of the participants conveying that had meaningful experiences. For example, in response to the individual quests, the most frequently coded subcategory under the general positive feelings category represents the opinion that a quest was meaningful or useful; 80% of the participants were coded to the subcategory at least once. Also in response to the individual quests, the third most frequent opinion was that a quest encouraged meaningful personal reflection; 90% of the participants mentioned it at least once. The participants said that they reflected about their own lives and their English ability; for example,

- In response to Quest 2, Participant 32 wrote: “I chose Quest 2 because I never think about my strengths and weaknesses. It was difficult for me to find out what my strengths and weaknesses are. After I finished the quest, I found out what I should do to improve myself.”
- In response to Quest 5, Participant 33 wrote: “I chose Quest 5 because I wanted to write about how my mother always mentally supports me. Recently, because I’m busy, I tend to forget to appreciate the kindness that other people show me and forget to give kindness to other people. This quest made me more aware. It’s good.”
- In response to Quest 5, Participant 38 wrote: “I chose Quest 5 because I didn’t know the idea of ‘Pay it forward’. I became interested in connecting acts of kindness in my daily life. Recently in my university club, one member said some wonderful and encouraging things to me. I wanted to share them with my classmate because they might have a good effect on them. It was difficult to express what I wanted to say in English, so I need to continue studying hard.”

When the participants reflected on the QBL experience, the data show that 22% of the participants said that it provided them with meaningful opportunities; it was the fifth most common perception of QBL. In her semi-structured interview, Participant 12 explained, “The questions that I answered during the quests were not questions I usually think about. For example, I had to explain why I study English. I don’t really talk or think about that myself, so it was a good opportunity to reflect on it”. Finally, when the participants shared their reflective opinions about the class in general, the fifth most frequent opinion, with 12 participants (28%) mentioning it, was that the class allowed them to have positive new experiences through completing the quests. The examples of quests fostering meaningful experiences provide evidence that positive psychology was successfully incorporated into the quests to support intrinsic FL motivation growth.

Most of the participants perceived the QBL experience to be suitably challenging. SDT explains that intrinsic motivation fosters through activities that are suitably challenging for the participant. The participants in this study were not specifically asked if the quests were challenging; however, the results suggest that the quests were suitably challenging. First, in response to the individual quests, 83% of the participants mentioned at least once that a quest was challenging but they could complete it in the end; it was the fifth most frequent opinion towards the individual quests. Second, when the participants reflected on their QBL experience, the seventh most common perception was that it was challenging but resulted in the participant learning something. Third, in the semi-structured interviews, 55% of the participants said that the quests were initially difficult but became easier. Fourth, when the participants shared their reflective opinions about the class in general, the third most frequent opinion, with 19 participants (44%) mentioning it, was that the class was challenging but worthwhile. The four instances of challenge provide evidence that the QBL experience fostered intrinsic motivation by providing a suitable level of challenge that the participants were able to manage successfully. The finding supports Lambert et al. (2015) and Çakıroglu et al. (2017) who also found that quests that are suitably challenging foster intrinsic motivation.

The participants appreciated the autonomy that quest choice provided. SDT explains that intrinsic motivation fosters through activities that support autonomy. Choice is

synonymous with autonomy. Therefore, quest choice was targeted as an area of research interest due to its potential to effectively support intrinsic motivation by fostering a feeling of autonomy. Two aspects of quest were explored: why a quest was chosen, and who should choose which quest to work on.

After completing a quest level, the participants explained why they chose to do the quests they did. The results of the data analysis conducted on the 355 identified reasons show that the ten most frequent reasons for choosing to do a quest were: (1) the quest looked easy or quest to finish; (2) they wanted to improve or use their English skill in the manner stipulated in the quest; (3) the quest looked interesting, fun or enjoyable; (4) they wanted to reflect on as aspect of their life that the quest was targeting; (5) they wanted to learn about foreign countries and culture through the quest; (6) they wanted to avoid making a video; (7) they wanted to do that type of quest; (8) they wanted to share something about themselves with the class through the quest; (9) they wanted to achieve a goal that the quest was encouraging; and (10) they wanted to try something new. This section discusses how the specific reasons for choosing to do a quest provide many examples of quest choice supporting and fostering intrinsic motivation.

The most frequent reason for choosing to do a quest was because it looked easy or quick to finish. The reason represents 13% of the data set; 18 participants mentioned it at least once. The finding that the participants most often choose quests because they looked easy or quick to finish directly supports Haskell (2012) who found the same result. On the surface, the finding suggests that the participants were not engaged in the learning process. However, from an SDT perspective, the finding hints that autonomy was supported, at least not negatively impacted, by not forcing the participants to do quests they did not want to do.

The second most frequent reason for choosing to do a quest was because the participant wanted to improve or use their English skills in the manner stipulated in the quest. The reason represents 13% of the data set; 23 participants mentioned it at least once. The reason suggests that quest choice supported intrinsic motivation by allowing the participants to choose quests that were suitable for their EFL

competence threshold.

Regarding the remaining reasons, according to the positive psychology literature (e.g., Seligman, 2008), the participants choosing a quest because it looked interesting, fun, or enjoyable, or choosing a quest because they wanted to reflect on their life suggests that quests have potential to increase intrinsic motivation through meaningful activities and positive emotions. The participants choosing a quest because they wanted to share something about themselves suggests that they are choosing quests which have the potential to foster intrinsic motivation through relatedness. The participants choosing a quest because they wanted to learn about foreign countries and cultures suggests that the quest has potential to foster intrinsic motivation through international posture (Yashima, 2002).

In the final quest diary, the participants shared their opinion about whether the teacher or the participant should choose which quest to do. The final quest diary data show that the participants overwhelming responded that it is better if the participant chooses which quest to work on; only three participants (7%) felt that the teacher should decide. The analysis of the 39 responses that were in favour of the participant deciding show four main reasons why: (1) 15 participants (36%) said it was because they can choose quests that are personally interesting; (2) 9 participants (21%) said it was because they could choose quests depending on their current life and study load; (3) 9 participants (21%) explained it was because being able to choose which quest to do makes them feel more motivated to complete it; and (4) 6 participants (14%) said it was because they could choose quests based on difficulty and quest requirements. For example,

- Participant 28 wrote: “It was very good that I could choose. I could work on quests which I was interested in. It would be difficult if the teacher chose the quest because everyone has quests they can work on or not depending on their environment.”
- Participant 31 wrote: “Of course I liked that I could choose because I think assigned homework by teacher is for high school not for University. It is important to work on study by own choice.”

The participants' preference for being able to choose which quest to work on was confirmed in the semi-structured interviews in which all the participants agreed with the sentiment. The reasons for wanting to choose which quest to work on suggest that the autonomy fostered through quest choice supports intrinsic motivation by allowing participants to choose personally interesting quests, and quests that are suitable for their current competence threshold. The finding supports De Schutter and Abeele (2014) who also found that students strongly agreed that they should be able to choose which quest to work on, and choosing quests was motivating, enjoyable, and engaging.

The QBL implementation encouraged the participants to have meaningful interpersonal experiences. SDT explains that intrinsic motivation fosters through activities that facilitate relatedness. Therefore, collaboration was targeted as an area of research interest due to its potential to effectively support intrinsic motivation by fostering a feeling of relatedness. The design of the QBL system in this study attempted to foster relatedness in two ways: first, through specific quests that required collaboration; and, second, through the quest scoring system that relied on peer assessment that forced the participants to see other participants' completed quests and comment on them. The results of the data analysis provide various examples of the quest fostering relatedness.

First, in the final quest diary, the participants wrote their opinions about the collaboration aspect of the questing experience. The data show that most of the participants had positive perceptions towards the collaboration aspect of the QBL experience. This is important because if the participants did not enjoy the collaboration aspect, this could negatively affect their intrinsic motivation; perceived forced collaboration could lead to feelings of being controlled, which is the opposite of intrinsic motivation according to SDT (Ryan & Deci, 2009). The participants explained that the collaboration aspect of QBL was enjoyable, provided good opportunities to interact with the other participants, strengthened existing relationships, and allowed the them to make new friends. Some examples that highlight the participants' positive perceptions of the collaboration aspect of QBL include:

- Participant 1 wrote: “For one of the quests, my friend and I interviewed each other about the jobs we want to do in the future. It was a good experience because it was the first time to actually record myself speaking in English. I also found out that my friend is thinking seriously about her future, so it made me motivated.”
- Participant 37 wrote: “I didn’t really like writing in my blog, but it was good that I did because I could look back at all my completed quests. I felt excitement when I looked at others’ blogs, I felt motivated. It was very good that I could see how they completed their quests. I was able to learn various ways of thinking and new English expressions.”
- Participant 39 wrote: “The collaboration was the best part of the homework.”
- Participant 44 wrote: “I could enjoy the quests more when working with other people rather than doing by myself.”
- Participant 46 wrote: “I think the collaboration of the quests is good because I can get a chance to communicate with my classmates in English.”

Second, the participants’ opinions about the individual quests they completed show that 65% of the participants freely mentioned at least once they had an interpersonally positive experience; it was the seventh most frequent opinion towards the individual quests. The participants explained that they enjoyed working with other people to complete the quest, enjoyed sharing their life with other classmates, and they enjoyed learning about their classmates. Third, when the participants reflected on the QBL experience, 26% of the participants specifically said that they appreciated the social aspect; it was the third most common reflection of QBL. The participants specifically mentioned that they liked learning about and from the other participants, and they liked writing and sharing their weekly blog. Fourth, the results of the semi-structured interviews show that 5 participants (55%) said they enjoyed the social aspect of questing. Fifth, when the participants shared their reflective opinions about the class in general, 9 participants (21%) mentioned that they could make friends and develop existing relationships.

Most of the participants clearly enjoyed the collaboration aspect of QBL. The finding supports (Lounis et al., 2014) who also found a preference in gamified contexts for collaboration on activities rather than doing them individually. This is important because in gamified environments, collaboration can increase performance and motivation (McGonigal, 2011; Ryan, Rigby, & Przybylski, 2006; Sheldon, 2011). In an EFL context, collaboration can reduce communication-related anxiety, which leads to increased self-confidence and motivation (Koga, 2010). The results related to collaboration clearly show multiple instance of the quests fostering meaningful relatedness. The finding that the quests fostered relatedness support Lambert et al. (2015) who also found that QBL was able to foster relatedness more than a traditional classroom experience, which lead to higher levels of enjoyment.

5.3.3 How QBL Hinders Intrinsic Motivation and FL Acquisition.

The analysis of the participants' opinions and perceptions showed generally positive opinions and perceptions of QBL; however, not all the data were positive. First, in response to the individual quests, the data show that a negative feelings category of response emerged from the data; 17 participants (42%) were coded to the category at least once. The negative feelings category represents only 3% of the opinions towards the individual quests. Second, when the participants reflected on the aspects of the QBL experience that they liked and did not like, 28 participants (67%) provided a response that was categorised as negative; 95% of the participants provided a positive response. It is important to note that the participants were encouraged to write about the things they liked and the things they did not like; a negative or positive reflection does not necessarily reflect a strong opinion. Finally, when the participants reflected on the quest choice and collaboration components of QBL, further negative issues emerged from the data. This section discusses how closer analysis of the data reveal three aspects of QBL that could hinder intrinsic motivation and FL acquisition: (1) difficulty of the quests; (2) time required to complete quests; and (3) quest choice allowing participants to avoid certain quests.

Some aspects of QBL were difficult. Regarding the negative feelings category of

response that emerged from the participants opinions of the quests they completed, the data show that the most frequent reason was because it was difficult. 11 participants (27%) were coded to the subcategory at least once. The participants explained that a quest was either difficult for their English ability, difficult to achieve the goal of the quest, or technically difficult. When the participants reflected on the aspects of the QBL experience that they liked and did not like, 12 participants (29%) also said that the quests were technically difficult.

Quest-based learning was time consuming for some of the participants. The second most frequent response of the negative feelings subcategory represents the opinion that a quest was time consuming. Seven participants (17%) were coded to the subcategory at least once. When the participants reflected on the aspects of the QBL experience that they liked and did not like, 11 participants (26%) also said that the quests were time consuming. For example, Participant 22 said "I didn't like the parts that took a long time, for example, it was a little troublesome when needing to record a video. But there were many good points also like the opportunity to learn about podcasts or other tools which I don't usually use."

According to SDT (Ryan & Deci, 2009), if a participant is overwhelmed by the challenge, or finds some aspect unnecessarily troublesome, it is potentially demotivating. The finding that some aspects of QBL were difficult and time consuming suggests that the conclusion presented in Section 5.3.2, QBL supports intrinsic motivation because it is suitably challenging, is not true for all the participants. The concern about unsuitable challenge being demotivating adds to Çakıroglu et al.'s (2017) finding that a large number of quests to complete is demotivating for some participants. The finding supports Kapp, Blair, and Mesch's (2014) suggestion that gamification is only one solution to solve a problem, and there might be other solutions that are more effective.

Quest choice can potentially limit the ability of quests to foster intrinsic motivation. Even though quest choice was generally seen as a positive aspect of QBL, the data show that the participants often chose quests that looked easy or quick to finish. If the participants chose quests based on how easy they looked, they could have inadvertently avoided quests that were more suitable for their competence level, or

quests that fostered deeper feelings of relatedness if they perceived working with other people to be troublesome. The results from Haskell (2012) provide evidence to support the argument as he showed that even though students were more likely to choose simple task-based quests because they looked easy or quick to finish, they enjoyed the challenging goal-based quests more, possibly due to them fostering stronger feelings of competence and autonomy.

The worry that quest choice could lead to the participants avoiding certain types of quests was shown to be a legitimate concern in the data about collaboration. When the participants reflected on the collaboration aspect of the QBL experience, the data show that 11 participants (27%) said that they did not collaborate. Three reasons were provided about why they did not collaborate: (1) 5 participants said they just always chose quests they could do themselves; (2) 4 participants said that they were too shy or too busy to work with other people; and (3) 2 participants said that they could not organise their schedule to work with other people. For example, Participant 28 explained,

> "I didn't work on quests that needed cooperation because I was busy with my university club activity. I couldn't work afterschool with anyone, and it seemed hard to organise a free period with the other students. I am not good at asking people to work with me on things like this, so I thought it was hard." (Participant 28)

Many of the quests were designed to foster intrinsic motivation through relatedness. The finding that many of the participants avoided quests that required collaboration shows how quest choice can negatively affect intrinsic motivation. The pedagogical implications of this finding are further discussed in Section 6.2.2.

5.4 The Effect of Leaderboards and Quest-Based Learning on Foreign Language Motivation

This section discusses how the leaderboard and the QBL experience affected the participants' FL motivation. The discussion centres around the results of the LLOS; however, data collected from the leaderboard questionnaire, quest-diaries, and semi-structured interviews are introduced to support the discussion. The discussion in Section 5.4.1 focuses on how the leaderboard affected the participants' FL motivation. The discussion in Section 5.4.2 focuses on how the QBL experience affected the participants' FL motivation.

5.4.1 The Effect of Leaderboards on Foreign Language Motivation

RQ3 examines the effect leaderboards have on FL motivation due to conflicting literature that requires further exploration. On one side of the argument, there is concern that extrinsically driven leaderboards negatively impact students' intrinsic motivation; the literature (e.g., Hanus & Fox, 2015; Philpott, 2015a) warns that intrinsic motivation could be affected if the participants perceived a leaderboard as an instrument of control. On the other side of the argument is the literature (e.g., Çakıroglu et al., 2017; Gåsland, 2011; Mekler et al., 2013; O'Donnell et al., 2013) that asserts that leaderboards positively impact intrinsic motivation if the leaderboard supports SDT. In the middle of the argument is Richet et al. (2014) who explains that "combining a leaderboard with points adds a social dimension with an unknown effect on motivation: it may either promote intrinsic motivation by experiencing competence, or reduce intrinsic motivation, if perceived as controlling" (p. 37).

The LLOS data show that the combined extrinsic motivation subscale pre-test scores remained basically the same as the post-test scores for both classes, and the combined intrinsic motivation pre and post-test subscale scores increased by about the same amount for both classes. Based on those broad descriptive variables, it could be suggested that the leaderboard did not affect either extrinsic or intrinsic FL motivation. This section discusses how closer analysis of the individual subscale

scores suggest that even though there is truth to all sides of the how-leaderboards-affect-motivation argument, there is an alternative hypothesis that has not been addressed in the literature: leaderboards increase extrinsic motivation and hinder intrinsic motivation more than they support intrinsic motivation.

The results of the paired samples *t*-tests conducted on the three external subscales show no statistical differences between the pre- and post-test data for either Class 1 or Class 2; however, the descriptive statistics provide insight. The mean scores for the external and the identified regulation subscales are most insightful. For external regulation, the mean score increased for Class 1 (+.31), but slightly decreased for Class 2 (-.10). As external regulation is the strongest form of extrinsic motivation, the results suggest that the use of the leaderboard in Class 1 led to an increase in the participants extrinsic motivation.

SDT explains that extrinsic motivation is externally derived and controlling in nature. Data collected from the leaderboard questionnaire and semi-structured interviews (discussed in Section 5.2) provide evidence that the rise in external regulation can be attributed to the leaderboard using points and leaderboard rank to control behaviour. This excerpt from the semi-structured interview with Participant 2 illustrates the feeling of being controlled by the leaderboard:

- Interviewer: How did you feel when you checked your leaderboard ranking each week?
- Participant 2: I thought it will make everyone do their homework correctly.
- Interviewer: You too?
- Participant 2: Well…because it publicly displays who did and didn't do their homework, if I didn't do my homework everyone would find out.
- Interviewer: I understand. You had a week in which you couldn't get 100 points right? How did you feel at that moment?
- Participant 2: I thought I should have completed my homework properly.
- Interviewer: So, this was the first time you had used a leaderboard in class? Did your feeling towards the leaderboard change as the semester progressed?
- Participant 2: At first, it was a new thing, so it was interesting. However, gradually I started to feel that I have to do my homework because the leaderboard will display whether I did it or not.

- Interviewer: You felt pressure?
- Participant 2: Yes, pressure.
- Interviewer: Did you talk about the leaderboard with your classmates? What did they think?
- Participant 2: Yes, they also said that because the leaderboard publicly shows who did and didn't do their homework, they felt like they should do their homework too.
- Interviewer: I understand. So, do you think using a leaderboard in class is a good idea?
- Participant 2: Yeah, I think so because it made everyone do more homework than they would have done if there was no leaderboard.

Various literature (e.g., Bielik, 2012; Hanus & Fox, 2015; Philpott, 2015a; Richet et al.; Tang & Hall, 1995) warns that a rise in extrinsic motivation comes at the cost of a decline in intrinsic motivation. The results of the paired samples *t*-tests conducted on the three intrinsic subscales show that the mean scores for Class 1 and Class 2 all increased. Therefore, the simple main finding could be that the use of a leaderboard did not negatively affect the participants' intrinsic FL motivation. However, closer analysis of all the LLOS data provide mixed evidence about the effect leaderboards have on intrinsic FL motivation. This section discusses three main findings that emerge from the data. The first and strongest finding is that leaderboards shift internally-leaning extrinsic motivation to externally-focused extrinsic motivation. The second finding is that leaderboards hinder the growth of intrinsic motivation represented by the knowledge subscale. The third and weakest finding is that leaderboards support the growth of intrinsic motivation represented by the accomplishment subscale.

The data suggest that the rise in external regulation for Class 1 came at the cost of identified regulation. The LLOS data show that identified regulation declined for Class 1 (-.44), but remained the same for Class 2. Identified regulation is a somewhat internally-leaning extrinsic motivation that fosters when external goals become personally important. If the participants in Class 1 were constantly exposed to a leaderboard that they considered not personally important, this could have resulted in the decline of identified regulation.

The leaderboard questionnaire and semi-structured interview data show that a small number of the participants provided comments that suggest they did not perceive the leaderboard as personally important; they said that they did not care about it, and it did not represent learning. However, the small number of comments do not provide strong supporting evidence that the decline in identified regulation was due to the participants not caring about the leaderboard. It seems more likely that rather than the leaderboard directly impacting identified regulation, the leaderboard's impact on external regulation resulted in the participants' FL motivation shifting from identified regulation to external regulation in order to achieve the extrinsic goals of the leaderboard. Simply put, as identified regulation is intrinsically-leaning, the leaderboard encouraged FL motivation to trend towards external motivation, away from intrinsic motivation.

The data collected from the intrinsic motivation subscales are mixed. The data suggest that the leaderboard hindered knowledge, supported accomplishment, and did not impact stimulation. For the knowledge subscale, the mean score rose only slightly for Class 1 (+.02), but a statistically significant rise was recorded for Class 2 (+.40). The result suggests that the presence of the leaderboard in Class 1 hindered the growth of the participants' intrinsic motivation that is derived from the pleasure and satisfaction of learning something new.

For Class 1, the rise in external regulation at the cost of identified regulation, and the lack of increase for the knowledge subscale show two ways the leaderboard negatively impacted intrinsic FL motivation. The findings support the literature (e.g., Bielik, 2012; Hanus & Fox, 2015; Philpott, 2015a; Richet et al.; Tang & Hall, 1995) that warns that a rise in extrinsic motivation negatively affects intrinsic motivation. Also, the findings not only support Deci et al. (1999) who found that extrinsic rewards that are performance-contingent undermine intrinsic motivation, but further show that they specifically undermine the intrinsic FL motivation measured by the knowledge subscale.

For the accomplishment subscale, even though no statistically significant increases were reported, the descriptive statistics show that the mean score increased more for Class 1 (+.36) than Class 2 (+.10). The accomplishment subscale represents

motivation that fosters through a feeling of satisfaction when a personally-meaningful L2/FL goal is achieved. Therefore, as the leaderboards rewarded weekly FL goals, this could have allowed the accomplishment type of intrinsic FL motivation to foster more for the participants in Class 1 than Class 2. Data collected from the leaderboard questionnaires and semi-structured interviews support the argument. For example, when the participants looked at the leaderboard they often self-reported the emotions determined and enjoyment. The emotions determined and enjoyment suggest psychological well-being which according to SDT and the positive psychology literature (e.g., Seligman, 2008), lead to, and are signs of intrinsic motivation towards the domain. The participants explained that they chose those emotions because they were motivated by rank, the score system was motivating and provided useful feedback, and they felt a sense of accomplishment.

The positive comments about the score system, and the sense of accomplishment bode well for intrinsic motivation. The literature (e.g., Kapp, 2012; Ryan & Deci, 2009) explains that if the participants feel that extrinsic rewards, such as points, are informative rather than controlling, this supports intrinsic motivation. Informative rewards could support autonomy through feedback. The participants explaining that they choose those emotions because they felt a sense of accomplishment appears to directly relate to the accomplishment subscale and suggests that the leaderboard could have led to the greater increase for the intrinsic motivation accomplishment subscale through a feeling of competence the participants felt when they viewed their score and rank on the leaderboard each week.

The data lightly suggest that leaderboards support the type of intrinsic FL motivation measured by the accomplishment subscale. The finding provides light support to the literature (e.g., Aldemir et al., 2018; Çakıroglu et al., 2017; Gåsland, 2011; Mekler et al., 2013) that argues leaderboards support intrinsic motivation because they foster a feeling of competence through informative feedback. The finding only provides light support because the increase for Class 1 was not statistically significant, and it was not significantly larger than the increase for Class 2. The finding differs from the literature (e.g., O'Donnell et al., 2013; Sheldon, 2011) that asserts leaderboards support intrinsic motivation through a feeling of relatedness that comes from the shared community created by a leaderboard. The data did suggest that the social

aspect increased extrinsic motivation, but the data did not suggest that the social aspect of the leaderboard was intrinsically motivating.

5.4.2 The Effect of Quest-Based Learning on Foreign Language Motivation

RQ3 examines the effect QBL has on FL motivation to determine its viability to be an intrinsically motivating approach for EFL pedagogy. The analysis of the LLOS external motivation subscales did not suggest that QBL affected extrinsic FL motivation in anyway. The finding was expected as the quests were specifically designed to foster intrinsic FL motivation. However, the analysis of the intrinsic motivation subscales strongly suggests that the QBL successfully increased the participants' intrinsic FL motivation. For example, the results of the paired samples *t*-tests conducted on the combined intrinsic motivation pre- and post-test mean scores reported a statistically significant rise in mean score for Class 1&2 (+.27); Class 2 (+.31) recorded a slightly higher rise than Class 1 (+.28). This section discusses the results of the individual intrinsic motivation subscales to show what types of intrinsic FL motivation the QBL experience supported.

The stimulation subscale represents intrinsic motivation derived from the positive feelings associated with doing an activity in the L2 (Vallerand & Ratelle, 2002). The results of the paired samples *t*-tests conducted on the stimulation pre- and post-test mean scores reported a statistically significant rise in mean score for Class 1&2 (+.38); Class 2 (+.43) recorded a slightly higher rise than Class 1 (+.33). For all the seven subscales, the pre and post-test mean scores changed the most for the stimulation subscale. The quest diary data discussed in Section 5.3 provide evidence that the participants enjoyed the quests and enjoyed using English to complete the quests. Participant 43's reflection of Quest 5 exemplifies the positive feeling of doing an activity using English:

> "I chose Quest 10 because it seemed not difficult and I like watching English videos. Watching English news in Quest10 was very good for studying English. It was good listening practice and I could learn news, so it was like killing two birds with one stone." (Participant 43)

The knowledge subscale represents intrinsic motivation derived from the pleasure and satisfaction of learning something new in the L2 (Vallerand & Ratelle, 2002). The results of the paired samples *t*-tests conducted on the knowledge pre- and post-test mean scores reported a statistically significant rise in mean score for Class 2 (+.40); the mean score for Class 1 (+.02) only slightly increased. Section 5.4.1 explained that the greater increase in the knowledge mean score for Class 2, along with the increase for Class 1 for the external regulation subscale, could suggest that by not applying leaderboard pressure on the participants in Class 2, it allowed them to focus more on the learning experience, leading to the greater increase in the intrinsic motivation accomplishment subscale. The leaderboard could have increased the external control and reduced the autonomy of the participants in Class 1, which affected the ability of the quests to increase the intrinsic motivation measured by the knowledge subscale. The quest diary data discussed in Section 5.3 provided evidence that shows that while completing quests, the participants derived pleasure and satisfaction from using their EFL skills to learn meaningful knowledge. Participant 3's reflection of Quest 14 exemplifies the positive feeling of learning using English:

> "I chose Quest 14 because I wanted to listen to people from various countries share their opinions about one topic in English. I could learn many things I don't usually learn. The ELLLO website has various topics so I think I can learn many things I'm interested in using English." (Participant 3)

The accomplishment subscale represents intrinsic motivation derived from achieving a personally-meaningful L2-related goal (Vallerand & Ratelle, 2002). The results of the paired samples *t*-tests conducted on the accomplishment pre- and post- test mean scores did not report any statistically significant changes; however, the descriptive statistics show that the mean scores increased for all three groups: Class 1 (+.36), Class 2 (+.10), Class 1&2 (+.24). The small increases in the mean scores for the accomplishment subscale could lightly suggests that the QBL experience increased the participants' intrinsic FL motivation related to achieving L2-related goals. The quest diary data discussed in Section 5.3 provide evidence that the quests successfully fostered feelings of competence and accomplishment. Participant 31's reflection of Quest 11 exemplifies the challenge and accomplishment feeling:

> "For this quest, I interviewed a foreign exchange student. At first, I thought it would be hard for a Japanese person to interview an exchange student only in English. I can't communicate well and wasn't sure if my English would be ok. So, I prepared well, and the interview went smoothly. I think I could learn the manner to communicate with exchange students." (Participant 31)

The LLOS data provide strong evidence that the QBL implementation successfully increased the participants' intrinsic FL motivation. Section 5.3.2 suggested that the QBL experience was intrinsically motivating because it was enjoyable, it was personally meaningful, it fostered competence through challenge, it fostered autonomy through quest choice, and it fostered relatedness through collaboration. The across the board rise of the intrinsic motivation subscale mean scores in conjunction with the reasons why the participants perceived the QBL experience to be intrinsically motivating, provide triangulated evidence that the QBL implementation increased the participants' intrinsic FL motivation. The finding supports SDT literature (e.g., McEown et al., 2014; Noels et al., 2001; Noels, 2005; Reeve, 2002; Vallerand, 1983; Zhou, Ma, & Deci, 2009) that explains how environments that foster competence, autonomy, and relatedness lead to an increase in self-determined intrinsic motivation. The finding not only supports Sheldon's (2012) assertion that the use of quests in a gamified course positively affect motivation, but more specifically presents reasons how it can occur in a gamified EFL course.

5.5 Summary

This chapter has presented a range of findings that highlight how leaderboards and QBL affect performance and FL motivation in a gamified EFL course. Section 5.2 discussed that the participants in Class 1 generally felt positively about the leaderboard. The participants explained that the points system motivated them by providing goals, and feedback about their performance. Closer analysis of the performance-related data revealed that the point system limited performance as it only encouraged performance up the maximum point threshold. This led to the

participants in Class 2 completing more homework than the participants in Class 1 as they were not constantly to the weekly point goals.

A participant's leaderboard rank was seen to have a noticeable impact on performance. The data showed that the participants who were low ranked on the leaderboard felt more negative emotions and attitudes towards the leaderboard, which in turn led to low performance scores. The low ranked participants explained that the negative emotions stemmed from their negative feelings about their rank and the public display of their rank. The results suggested that the low ranked participants were initially engaged, but their engagement declined as their rank declined. In contrast to the low ranked participants, the higher ranked participants reported more positive emotions and attitudes towards the leaderboards which corresponded with more successful performance scores for those participants. The participants explained that the positive emotions stemmed from feeling motivated to gain or maintain rank, responding positively towards the point scoring system, and feeling a sense of accomplishment associated with their leaderboard standing.

Section 5.3 reported that the participants had overwhelming positive opinions and perceptions towards the QBL experience. The participants perceived QBL as an effective approach for EFL pedagogy as it provided them with many opportunities to use or improve their English skills, showed them new ways to study English, and encouraged them to use their English skills in meaningful situations. The participants also perceived QBL to be an intrinsically motivating approach for EFL pedagogy as it was enjoyable, personally meaningful, fostered competence through challenge, autonomy through quest choice, and relatedness through collaboration. Some concerning results were reported about the quests being difficult and time consuming, participants often choosing quests based on how easy they look, and some participants avoiding quests that required collaboration. However, the concerning results were outweighed by the positive results.

Section 5.4 reported the effect leaderboards and the quests had on FL motivation. The section discussed how the results of the LLOS showed that leaderboards increased external regulation, decreased identified regulation, and hindered the

knowledge type of intrinsic motivation. Light evidence was presented to suggest that the leaderboard fostered the type of intrinsic motivation represented by the accomplishment subscale. However, the data most strongly suggested that the leaderboard increased extrinsic motivation, and hindered intrinsic motivation more than it supported intrinsic motivation. The section then discussed how the results of the LLOS provided strong evidence that the quests led to an increase in the participants intrinsic FL motivation; the mean scores increased for all the intrinsic motivation subscales, with the increases being statistically significant for Class 2 for the knowledge subscale, Class 1&2 for the stimulation subscale, and Class 1&2 for the combined intrinsic motivation subscales.

Chapter 6. Conclusions and Implications

6.1 Conclusions

The classroom-based research presented in this thesis explored the use of leaderboards and quests in a gamified EFL course at a Japanese university. This chapter provides a summary of the findings and discusses their theoretical, pedagogical, and methodological implications. The chapter concludes the thesis by presenting the limitations of the study with suggestions for future research.

6.1.1 The Effect of Leaderboards on Performance

RQ1 sought to determine the effect leaderboards have on student performance. The leaderboard questionnaire data showed that most of the participants in Class 1 responded positively to the leaderboard. When they looked at the leaderboard, they reported the positive emotions enjoyment, hope, pride, and determined noticeably more often than the negative emotions anxiety, shame, hopelessness, and envy. The participants explained that they chose the positive emotions because they were motivated by the ranking system, they were responding positively to an aspect of the point system, and they felt a sense of accomplishment when they looked at the leaderboard. When the participants chose a positive emotion, they were more likely to achieve a full performance score in the following week compared to the participants who chose a negative emotion.

Even though the participants responded positively to the leaderboard, the performance-related data showed that the class without the leaderboard, Class 2, recorded higher mean scores for all measures of performance; slightly higher for the English Central and quests completed measures, and noticeably higher for the MReader performance measures. Class 2 basically achieved all five of their final target point totals, whereas, Class 1 only achieved four of their final target point totals. The data showed a larger group of high performing participants in Class 2 compared to Class 1 who pushed up the performance scores for Class 2. The identification of the group was significant as it led to four main findings about how leaderboards affect performance in a gamified EFL course.

The first finding about leaderboards is that the point system encourages performance but also limits performance. The performance-related data showed that most of the participants in Class 1 did enough work to achieve their final target point totals, but not more. In contrast, a larger group of high performing participants in Class 2 completed noticeably more work than the final target point totals. The leaderboard questionnaire data showed that when the participants self-reported positive emotions to represent how they felt about the leaderboard, they often explained that it was because the point system was a motivating force that provided clear goals, feedback, and contributed to a feeling of accomplishment. The data suggest that for Class 1, the constant exposure to the leaderboard and point system increased cognisance of the weekly performance goals which led to the participants achieving their weekly goals. However, once the goal was achieved and no more rewards could be received, performance stopped. As the participants in Class 2 were not constantly exposed to the point system, they did not think about the weekly performance goals as much, which led to a larger amount of participants completing more work than required by the point system.

The finding about the point system aspect of the leaderboard supports the literature (e.g., Aldemir et al., 2018; Domínguez et al., 2013; Iosup & Epema, 2014; Tan & Hew, 2016) that found gamification components, such as point systems, encourage performance by providing clear goals and feedback. However, the finding more strongly supports the literature (e.g., Borys & Laskowsky, 2013; Kapp, 2012; Koivisto & Hamari, 2014; Krause, Mogalle, Pohl, & Williams, 2015; Mollick & Rothbard, 2014; Sheldon, 2011) that argues extrinsic goals create a performance ceiling and once the goal has been achieved, the performance is unlikely to continue if not rewarded. The finding about the point system is significant because it shows how a point system can be beneficial for situations that have specific goals; however, if the goal is maximum performance, a point system that has a maximum reward could limit performance.

The second finding about leaderboards is that they could negatively impact the quality of performance if the participants feel that achieving the extrinsic goals, such as high leaderboard rank or maximum points, is more important than having

meaningful learning experiences. The leaderboard questionnaire and semi-structured interview data showed a small number of participants commented that the point system should have better reflected the quality of work, not just quantity. The number of comments were small; however, it does not mean it was not true for other participants. It was possible that some of the participants mindlessly completed the learning activities in order to get points. The comments provide light support for the literature (e.g., Domínguez et al., 2013; Halan, Rossen, Cendan, & Lok, 2010; Tan & Hew, 2016; Werbach & Hunter, 2012) that asserts, if a point system prioritises quantity over quality, it may negatively affect the quality of the learning outcomes. The comments also support the warnings (e.g., Aldemir et al., 2018; Hanus & Fox, 2015) that extrinsic rewards, such as points, need to reward the desired behaviour, not just behaviour.

The third finding about leaderboards is that they positively affect the emotions, attitudes, and performance of the participants who have a middle or high leaderboard rank. The middle and high ranked participants reported the positive emotions determined, enjoyment, hope, and pride more frequently than the low ranked participants. They explained that they reported the positive emotions because they were motivated to maintain or gain rank, they were motivated by the score system because it tracked their progress, and they felt a sense of accomplishment when they looked at the leaderboard. Those participants who reported the positive emotions were more likely to achieve a full score in the following week compared to the participants who reported negative emotions. Compared to the low ranked participants, those higher ranked participants were motivated to maintain their rank, thought the leaderboard was fun, and had stronger positive attitudes towards the learning activities.

The finding that leaderboards positively affect higher ranked participants supports the literature (e.g., Aldemir et al., 2018; Buckley, Doyle, & Doyle, 2017; Çakıroglu et al., 2017; Heeter, Lee, Medler, & Magerko, 2011; Tan & Hew 2016; Tran & Zeckhauser 2012; Werbach & Hunter, 2012) that explains that a competitive social leaderboard positively affects the performance of higher performing students. The finding that leaderboards positively affect students with a high rank is significant

because it provides insight into future gamified instruction design. For example, if all students could be encouraged to believe that they have a good leaderboard rank, it could sustain performance and positive attitudes towards the learning activities.

The fourth finding about leaderboards is that they negatively affect the emotions, attitudes, and performance of students who have a low leaderboard rank. The low ranked participants reported the negative emotions anxiety, shame, and hopelessness more frequently than the higher ranked participants. Those low rank participants explained that they reported the negative emotions due to their negative feelings about their rank, their score, and the forced social comparison created by the leaderboard. Those participants who reported the negative emotions were less likely to achieve a full score in the following week compared to the participants who reported the positive emotions. Compared to the higher ranked participants, those low ranked participants were not motivated to maintain rank, did not think the leaderboard was fun, and had stronger negative attitudes towards the learning activities.

The low ranked participants initially reported positive emotions, attitudes, and performance. However, as the semester progressed and their rank failed to rise, the positive results turned to negative results. The decline in sentiment and performance partially supports the literature (e.g., Borys & Laskowsky, 2013; Buckley et al., 2017; Garland, 2015; Koivisto & Hamari, 2014; Krause et al., 2015; Mollick & Rothbard, 2014; Werbach & Hunter, 2012) that warns gamification techniques, such as leaderboards, are good for short-term performance boosts but will not be able to sustain engagement over an extended period-of-time once the novelty wears off. This finding only partially supports the literature as the decline in performance was only registered for those low ranked participants.

The literature (e.g., Barata et al., 2013; Charles et al., 2011; Domínguez et al., 2013; Nicholson, 2013) argues that the competitive and comparative aspect of a leaderboard in an education setting negatively affects some students. The finding that the low ranked participants responded negatively to a leaderboard supports the literature (e.g., Aldemir et al., 2018; Buckley et al., 2017; Çakıroglu et al., 2017; Tan & Hew 2016; Werbach & Hunter, 2012) that warns leaderboard rank negatively

affects the performance of lower performing students. This finding is significant because prior to the study, the researcher assumed that the competitive social aspect of the leaderboard would encourage the low performing participants to work harder. However, the opposite occurred, the negative feelings led to negative performance which led to further negative feelings. This finding exemplifies Dörnyei and Ushioda's (2011) explanation of the cyclic nature of motivation and performance: positive performance fosters motivation which leads to positive performance. Further research is required to determine if the low performing students would have performed differently if there was no leaderboard.

6.1.2 Students' Opinions and Perceptions of Quest-Based Learning

RQ2 sought to explore students' opinions and perceptions of QBL to determine its viability as an intrinsically motivating approach to EFL pedagogy. The content analysis of the qualitative data collected from the quest diaries and the semi-structured interviews showed overwhelmingly positive opinions and perceptions towards the QBL experience. Three major themes emerged from the data: (1) how QBL supports FL acquisition; (2) how QBL supports intrinsic motivation; and (3) how QBL hinders intrinsic motivation and FL acquisition. The following paragraphs summarise the main points of the findings.

The first reason why the participants perceived QBL to be beneficial for FL acquisition was because it provided them with many opportunities to use or improve their English skills. The second reason was because they learned many new ways to study EFL. The third reason was because it allowed them to use English in meaningful situations to learn. The findings not only support the literature (e.g., Çakıroglu et al., 2017; Lambert, et al., 2015) that found students have positive perceptions of QBL, but further provides three reasons why students have positive perceptions of QBL as an approach for FL acquisition. The finding is significant as this is the first major investigation into the use of QBL for FL acquisition. It shows how TBLT (e.g., Ellis, 2003; Willis & Willis, 2007) can be combined with SLA theory (e.g., Krashen, 2009; Swain, 1985) to design quests that are suitable for FL acquisition.

The first reason why QBL was intrinsically motivating was because it was enjoyable. Positive psychology literature (e.g., Seligman, 2008) and SDT (Ryan & Deci, 2009) suggest that positive emotions, such as enjoyment, reflect psychological well-being which leads to intrinsic motivation. Many instances of the participants reporting enjoyment emerged from the data. The finding that the participants perceived QBL as an enjoyable experience supports Lambert et al. (2015) and further shows how the autonomous-nature of QBL leads to enjoyment if the participants can explore content they are interested in, without being forced.

The second reason why QBL was intrinsically motivating was because it provided personally meaningful experiences. Many instances of the participants reporting that the QBL experience was meaningful emerged from the data. Therefore, as the positive psychology literature (e.g., Seligman, 2008) explains that meaningful activities lead to intrinsic motivation, the participants' frequent comments reflecting a meaningful experience suggest that the QBL experience was intrinsically motivating.

The third reason why QBL was intrinsically motivating was because it fostered competence through challenge. Many instances of the participants reporting that the QBL experience was challenging emerged from the data. As SDT (Ryan & Deci, 2009) explains that activities that are suitably challenging are intrinsically motivating, the participants' frequent comments reflecting challenge suggest that the QBL experience was intrinsically motivating. The finding supports the literature (e.g., Çakıroglu et al., 2017; Lambert et al., 2015) that also found QBL can provide a suitably challenging experience to students. The design of the quests was influenced by TBLT literature (e.g., Ellis, 2003; Nunan, 1989, 1997) that states that the context, culture, sequencing, and grading of tasks requires consideration. The finding that the quests were suitably challenging supports designing EFL quests based on relevant findings from the TBLT literature.

The fourth reason why QBL was intrinsically motivating was because it fostered autonomy through quest choice. Quest choice was targeted as an area of research interest due to its apparent close alignment to the autonomy tenet of SDT. Two aspects of quest choice were explored: why a quest was chosen, and who should

choose which quest to work on. The reasons why a quest was chosen, and the large amount of reasons why the participant should choose which quest to work on provided strong evidence that the autonomy fostered through quest choice supports intrinsic motivation. The data further suggested that the autonomy lead to the participants choosing quests that foster intrinsic motivation through enjoyment, relatedness, international posture, and competence. The finding supports De Schutter and Abeele (2014) who also found that students strongly agreed that they should be able to choose which quest to work on, and choosing quests was motivating, enjoyable, and engaging.

The fifth reason why QBL was intrinsically motivating was because it fostered relatedness through collaboration. Collaboration was targeted as an area of research interest as QBL appeared to be uniquely positioned to provide many opportunities for meaningful interpersonal experiences. The results of the data analysis provided various examples of the quests fostering relatedness. The participants explained that the collaboration aspect of QBL was enjoyable, provided good opportunities to interact with the other participants, strengthened existing relationships, and allowed them to make new friends. The finding supports (Lounis et al., 2014) who also found a preference in gamified contexts for collaboration on activities rather than doing them individually. The finding that the QBL experience fostered relatedness supports Lambert et al. (2015) who also found that QBL was able to foster relatedness more than a traditional classroom experience. This finding is significant as it provides guidance for future quest design.

Even though the participants had generally positive opinions and perceptions towards the QBL experience, the data reveal three aspects of QBL that could hinder intrinsic motivation and FL acquisition. The first and second concern are about the quests undermining the competence tenet of SDT. First, some of the participants explained that a quest was either difficult for their English ability, difficult to achieve the goal of the quest, or technically difficult. Second, some of the participants explained that some of the quests were too time consuming. According to SDT (Ryan & Deci, 2009), if a participant is overwhelmed by the challenge, or finds some aspect unnecessarily troublesome, it is potentially demotivating. The concern about unsuitable challenge being demotivating adds to Çakıroglu et al.'s (2017) finding

that a large number of quests to complete is demotivating for some participants.

The third concern is about quest choice limiting the ability of quests to foster intrinsic motivation. Even though quest choice was generally seen as a positive aspect of QBL, the data show that the participants most often chose quests that looked easy or quick to finish. The finding directly supports Haskell (2012) who also found the same. This is a significant finding as it shows a paradox about fostering autonomy through quest choice. On one hand, allowing a choice of quest to complete supports the fostering of intrinsic motivation through autonomy. On the other hand, it leads to some students choosing quests that are simple to complete, rather than beneficial to complete. The worry that quest choice could lead to the participants avoiding certain types of quests was shown to be a legitimate concern in the data collected about collaboration. When the participants reflected on the collaboration aspect of the QBL experience, 11 participants (27%) said that they never chose quests that required collaboration.

6.1.3 The Effects of Leaderboards and Quests on Foreign Language Motivation

RQ3 sought to determine the effects of leaderboards and QBL on FL motivation. The LLOS was the primary instrument used for this investigation. However, all the data collection instruments provided important insight and led to three main findings: (1) leaderboards increase extrinsic FL motivation; (2) leaderboards undermine intrinsic FL motivation more than they support intrinsic motivation; (3) the quests increased intrinsic FL motivation. The following paragraphs summarise the main points of the findings.

The first main finding about the effect leaderboards have on FL motivation is that they increase extrinsic FL motivation. Even though the results of the paired samples *t*-tests conducted on the LLOS data reported no statistically significant changes for the extrinsic motivation subscales, the data clearly show that external regulation increased for Class 1 but declined for Class 2. As external regulation is the strongest form of extrinsic motivation, it shows that leaderboards increase extrinsic motivation, specifically, external regulation. External regulation fosters when an

external force uses rewards and punishment to control behaviour. The rise in external regulation was attributed to the rank and point system as they were often perceived as instruments of control.

The finding that external regulation increased is significant as it provides insight into how leaderboards affect performance and motivation. The rise in external regulation provides a warning that, once the rewards are taken away or lose their value, motivation to continue doing the learning activities the leaderboard rewards will disappear. The finding supports the literature (e.g., Borys & Laskowsky, 2013; Buckley, Doyle, & Doyle, 2017; Garland, 2015; Krause, Mogalle, Pohl, & Williams, 2015; Koivisto & Hamari, 2014; Mollick & Rothbard, 2014; Werbach & Hunter, 2012) that warns gamification techniques such as leaderboards will not be able to sustain engagement over an extended period-of-time once the novelty wears off. The finding is significant as most gamification literature until now has only been concerned with whether leaderboards positively or negatively affect intrinsic motivation.

The second main finding about the effect leaderboards have on FL motivation is that they undermine intrinsic motivation more than they support intrinsic motivation. Leaderboards negatively impact intrinsic FL motivation in two ways: (1) shift internally-leaning extrinsic motivation to externally grounded extrinsic motivation; (2) hinder the growth of intrinsic FL motivation that derives from the pleasure and satisfaction of learning something new in the L2/FL. Leaderboards possibly support the growth of intrinsic FL motivation that fosters through a feeling of satisfaction when a personally-meaningful L2/FL goal is achieved. The following paragraphs summarise the main points of the findings.

The data suggested that the rise in external regulation for Class 1 came at the cost of identified regulation; identified regulation declined for Class 1 but remained the same for Class 2. Identified regulation is a somewhat internally-leaning extrinsic motivation that fosters when external goals become personally important. The data suggested that rather than directly impacting identified regulation, the strong pull of external regulation shifts the internally-leaning extrinsic motivation to externally-focused extrinsic motivation. The emotion and attitude data collected from the

leaderboard questionnaire suggest that the decline in identified regulation could be limited to the low performing participants. Further research is required to determine if the suggestion is true. The finding that leaderboards increase external regulation at the cost of identified regulation makes a significant contribution to the gamified instructional design literature as it is the first enquiry into how leaderboards affect extrinsic FL motivation, most studies only consider the impact leaderboards have on intrinsic motivation. The decline in identified regulation and rise in external regulation provides support to the early critics (e.g., Bogost, 2011; Robertson, 2010; Sierra, 2011) of gamification who said that gamification replaces meaningful incentives with fictional incentives and it will not be able to sustain engagement once the novelty wears off.

The data suggested that leaderboards hinder the growth of intrinsic FL motivation that derives from the pleasure and satisfaction of learning something new in the L2/FL. The LLOS data showed that the knowledge subscale mean score basically remained unchanged for Class 1, but a statistically significant increase was reported for Class 2. The contrasting results suggest that the leaderboard could have hindered the growth of the intrinsic FL motivation measured by the knowledge subscale by encouraging the participants to focus on leaderboard rewards rather than the learning process. The finding supports Deci et al. (1999) who found that extrinsic rewards that are performance-contingent undermine intrinsic motivation. The lack of growth in the knowledge subscale, and the decline in the intrinsically leaning identified regulation provide support to the literature (e.g., Bielik, 2012; Hanus & Fox, 2015; Philpott, 2015a; Richet et al.; Tang & Hall, 1995) that warns a rise in extrinsic motivation comes at the cost of a decline in intrinsic motivation.

Leaderboards possibly support the growth of intrinsic FL motivation that fosters through a feeling of satisfaction when a personally-meaningful L2/FL goal is achieved. The LLOS data showed that the accomplishment subscale mean score increased more for Class 1 than Class 2. The increase was not statistically significant, and it was not massive, but it was noticeable. The literature (e.g., Kapp, 2012; Ryan & Deci, 2009) explains that, if rewards are informative rather than controlling, they can support the growth of intrinsic motivation. The leaderboard questionnaire data and semi-structured interview data showed that many of the participants commented

that the point system provided informational feedback that led to feelings of motivation and accomplishment. The finding provides light support for the literature (e.g., Çakıroglu et al., 2017; Gåsland, 2011; Mekler et al., 2013) that asserts leaderboards positively impact intrinsic motivation. However, as the finding about how leaderboards support intrinsic FL motivation is weaker than the two findings about how leaderboards negatively impact intrinsic FL motivation, the conclusion is that leaderboards hinder intrinsic motivation more than they support intrinsic motivation.

The main finding about the effect quests had on FL motivation is that they increased intrinsic FL motivation. The LLOS data showed that the mean scores for each of the intrinsic motivation subscales increased for both classes. The greatest increase was for the stimulation subscale. The paired samples *t*-tests conducted on the stimulation pre- and post-test mean scores reported a statistically significant rise in mean score for both classes combined. The accomplishment and knowledge subscale mean scores increased about the same amount for both classes combined. For the knowledge subscale, a statistically significant rise was reported for Class 2; however, the mean score remained about the same for Class 1. The greater increase in the mean score for Class 2, along with the increase in the mean score for Class 1 for the external regulation subscale, could suggest that, by not applying leaderboard pressure on the participants in Class 2, it allowed them to focus more on and enjoy the learning experience. The results of the paired samples *t*-tests conducted on the combined intrinsic motivation pre- and post-test mean scores reported a statistically significant rise in mean score for both classes combined.

The results of the LLOS clearly show that intrinsic FL motivation increased. The LLOS data suggest that the participants' intrinsic motivation grew when they used their English skills to complete quests, learn something, and achieve personally-meaningful English-related goals. Data from the quest diaries and semi-structured interviews triangulated the LLOS findings and provide further general suggestions that the quests were intrinsically motivating because they were enjoyable, meaningful, suitably challenging, supported autonomy, and interpersonally meaningful. The finding about QBL increasing intrinsic FL motivation is significant as it not only provides clear evidence supporting Sheldon's (2012) assertion that

QBL is a motivating pedagogical approach, but further shows how academic theory (e.g., SDT, positive psychology, TBLT) can be used to design intrinsically motivating quests in an EFL context.

6.1.4 Final Conclusions about Leaderboards and Quests

In conclusion, most of the participants, except the ones who appeared low on the leaderboard ladder, responded positively to the leaderboard. As the researcher had taught the participants in the previous semester, he suspects that without the leaderboard, on average, the participants in Class 1, even the low ranked ones, would have completed less work than they did. However, as the data collected in this study showed the participants in Class 2 completed more work than the participants in Class 1, it is hard to argue that the use of a leaderboard was more effective than just telling the participants what their weekly point goals were. What is clear is that the extrinsic rewards of points and rank are influential components of a leaderboard. Points and rank motivate behaviour by providing clear goals and feedback about performance, but points also limit performance, and the rank aspect is detrimental to the psychology of the low performing participants.

Even though points and rank can potentially support intrinsic FL motivation for high performing students, the more dominant effect is that they are perceived as instruments of external control that increase extrinsic FL motivation at the cost of intrinsic FL motivation. This provides three warnings. First, once the reward is taken, motivation to do the activity the leaderboard rewards will stop. Second, performance will decline. The low ranked participants performance will decline quicker as they are not receiving the positive feedback; however, eventually the higher ranked participants performance will decline as the inevitable hedonic adaption sets in (Chen, 2003; Ferster & Skinner, 1957). Third, if the impacted intrinsic motivation does not recover, the leaderboard could have a long-term negative impact on intrinsic FL motivation. Therefore, leaderboards are useful for environments that require a short-short term performance boost. Ideally, a leaderboard is used to encourage students to complete activities that increase their intrinsic motivation.

The findings about the leaderboard presented in this thesis are in relation to the specific type of leaderboard used in this study. The leaderboard was a limited point leaderboard as the participants could receive only up to 100 points each week. The researcher limited the weekly points in order to restrain competitive students taking overwhelming leads that demotivated other students because they could not catch up. However, while it successfully did this, it did not allow low-ranked students to catch up because the maximum weekly point total was 100. A system to allow the low-ranked students to catch up should have been considered. The leaderboard was completely public as all the participants could see each other's rank and score. A leaderboard that was configured differently, such as only showing the participant who was one rank higher could yield different results.

The quests successfully increased intrinsic FL motivation while supporting FL acquisition. This finding not only supports Sheldon's (2012) assertion that the use of quests in a gamified class positively affects motivation, but more specifically presents reasons how it can occur in a gamified EFL course. The positive results support the theoretical framework and suggest that the strength of the QBL system came from the combination of pedagogies and theories that were incorporated in the design. The findings about QBL are significant as they explore deeper and wider than any other enquiries into QBL, especially for the EFL context. The findings are significant for the Japanese context as it shows how QBL can be used to increase intrinsic FL motivation. The findings about the quests presented in this thesis are in relation to the specific quests used in this study.

6.2 Implications

This study aimed to provide knowledge to the education-related gamification research field by identifying specific aspects of leaderboards and QBL that affect performance and motivation. No other gamified instructional design studies have explored deeper and wider than this study. The findings from this study contribute to not only the educational-based gamification research field, but also the L2/FL motivation, general gamification, and mainstream psychology research fields.

6.2.1 Theoretical Implications

The results of the study contribute to the FL motivation literature by providing an updated perspective on the state of FL motivation at a Japanese university. The LLOS data showed that the participants in this study reported very low scores for the amotivation subscale. The results of the extrinsic motivation identified regulation subscale confirmed Yashima et al.'s (2009) finding that Japanese EFL learners strongly align with identified regulation. A strong alignment to identified regulation suggests that Japanese EFL students consider learning EFL an important task and they are willing to endure activities they do not intrinsically enjoy in order to improve their English ability. The participants also strongly aligned with the intrinsic motivation knowledge subscale; the strong alignment suggests that the students derive pleasure and satisfaction from learning something using English.

The low amotivation, high identified regulation, and high knowledge results reject Berwick and Ross's (1989) description of Japanese universities as *motivational wastelands*. As identified regulation was stronger than knowledge, the LLOS data provide support to the literature (e.g., Hamada & Kito, 2007; Ockert, 2011; Tachibana, Matsukawa, & Zhong, 1996) that found that Japanese students possess stronger extrinsic FL motivation than intrinsic FL motivation. However, as identified regulation is internally leaning and as the knowledge alignment was also strong, it does not suggest that there is a large gap between extrinsic motivation and intrinsic motivation. The results of the LLOS exemplify McEown et al.'s (2014) explanation that learners align with multiple types of motivation regulation.

The findings from the LLOS data contribute to the SDT literature by showing evidence that the OIT extrinsic motivation continuum is a valid construct and that leaderboards align with external regulation. The design of the gamification system was guided by the central belief of the OIT continuum that externally regulated behaviour can become internally integrated behaviour if the environment supports the psychological needs of autonomy, competence, and relatedness. However, as the opposite occurred, identified regulation declined and external regulation increased for Class 1, the LLOS data show how internally regulated behaviour becomes externally regulated behaviour if a psychological tenet is not supported. The results

of leaderboard questionnaire suggested that leaderboards could have decreased the feeling of autonomy and increased the feeling of control by rewarding and punishing behaviour using points and rank.

Japanese peoples' poor English ability is most often attributed to their low intrinsic FL motivation (e.g., McVeigh, 2004; Nakata, 2006; Ushioda, 2013). Identified regulation is classified as extrinsic motivation; however, it is considered an internally-leaning type of extrinsic motivation. The decline in identified regulation due to the leaderboard is significant for Japanese FL motivation as Japanese EFL learners align most strongly to it (e.g., Yashima et al., 2009). If the participants' external regulation remains high, and identified regulation remains low after the leaderboards are taken away, it could be considered a negative impact on long-term intrinsic FL motivation. The results of the study support the literature (e.g., Bielik, 2012; Hanus & Fox, 2015; Richet et al.; Tang & Hall, 1995) that warns a rise in extrinsic motivation comes at the cost of a decline in intrinsic motivation, with the caveat that it is the intrinsically-leaning identified regulation that declines.

Even though some L2/FL researchers are moving past SDT onto new frameworks (e.g., LTMSS, complex dynamic systems theory), the results of the study show SDT is still a useful theory to analyse FL motivation. The results of the study also show that SDT is particularly useful for educational-based gamification studies due to the range of extrinsic and intrinsic motivators implementations need to consider. Using the LLOS to analyse the participants' FL motivation was effective as it was able to clearly identify the participants' FL motivation orientations, align leaderboards to external regulation, show how the alignment to external regulation led to a decrease in identified regulation, and provide evidence that the quests increased the participants' intrinsic FL motivation.

SDT is a meta-theory of motivation that continues to evolve by incorporating new sub theories to further target specific aspects of extrinsic and intrinsic motivation. This study also incorporated other theories to strengthen the SDT foundation. Behaviourism was used to explain the ramifications of external rewards on performance in a more comprehensive way than the external regulation could. Various other cognitive theories (e.g., ideal-self, international posture, positive

psychology) were used to guide the design of intrinsically motivating quests, targeting relevant aspects not specifically accounted for in SDT. The successful inclusion of the psychological theories suggest ways that SDT could continue to evolve.

The results of the study contribute to the FL motivation, TBLT, and CALL research fields by introducing QBL as new framework to design and deliver intrinsically motivating learning materials. The analysis of the participants' opinions and perceptions towards QBL, alongside the increase of the three intrinsic motivation LLOS subscales provides evidence that the quests achieved their goal of increasing intrinsic FL motivation. The quests were successful because they leveraged various aspects of a number of theories and approaches (e.g., SDT, positive psychology, ideal-self, international posture). The CALL ecosystem connected everything together and allowed specific motivation constructs to be targeted. The QBL system provided a suitable level of linguistic and personal challenge, allowed autonomy to foster through quest choice, and facilitated situations for meaningful relatedness to occur. The results of the study provide new avenues for research to explore that combines aspect of FL motivation theory, TBLT, and CALL.

The study showed how a general gamification framework (e.g., Werbach & Hunter, 2012) can guide the selection of which gamification components to include in an implementation to achieve a specific goal. However, even though The Pyramid of Elements framework (Werbach & Hunter, 2012) provided a simple and robust framework, it did not include some aspects that were important for the application of gamification to an educational course. Chou's (2015) Octalysis framework provides more detail that should be considered for gamified instructional design implementations such as differentiating between extrinsic and intrinsic motivators, and differentiating between positive and negative motivators. Marczewski's (2015) Hexad player type framework draws heavily from SDT and provides a simple yet robust framework for gamification design; however, as it focuses on player types rather than process outcomes, it is not so suitable for gamified instructional design.

None of the frameworks are specifically designed for gamified instructional design. Even though the authors of the frameworks provide extra information accompanying

their frameworks to explain how their frameworks can be applied to any situation, their frameworks fail to account for issues that are pertinent to the educational context. The most obvious omission in the frameworks is that they do not consider the relationship between the learning theory and content, the gamification techniques, and the ecosystem. Therefore, based on the findings from this study, and borrowing from the three frameworks, the researcher has developed a new framework that is specifically designed for gamified instructional design. The new framework, entitled, The Gamified Instructional Design Framework (see Figure 6.1), has the simplicity of The Pyramid of Elements framework, the robustness of Hexad, and accounts for the range of psychological theories considered in Octalysis.

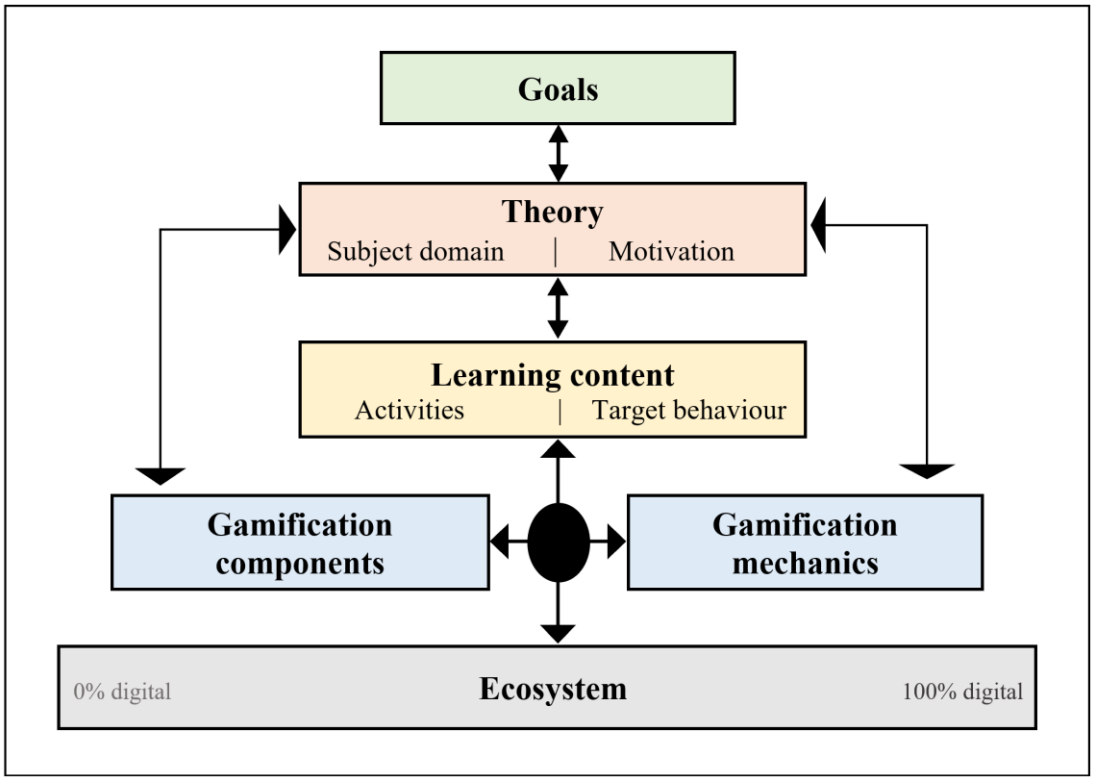


Figure 6.1. The gamified instructional design framework.

The gamified instructional design framework starts at the top with the goals of the implementation. After the goals have been determined, context-relevant theory should be considered to make sure the goals can be achieved. The theory should consider theory related to the subject domain, and theory that accounts for the effect of the gamification. Even though there is overlap about how students learn and

become motivated about different academic subjects, there are differences depending on course. For example, how students learn about history or maths will be different than how they learn a foreign language. The students' age and ability should also be considered.

The learning content is the centre of a gamified course. The learning content should be supported by the theory. The learning content not only includes the actual learning activities, but also considers how the students should interact with them. The target behaviour could aim to get students to focus on the amount or accuracy of completed work, learner autonomy, or meaningful collaboration. Then, gamification components and mechanics that support the learning content are chosen and calibrated.

The Ecosystem refers to the tools used to facilitate and connect the components, mechanics, and learning content. It includes the aesthetics of the gamified course. The ecosystem could be completely digital, partially digital, or completely non-digital. The double-ended arrows show that each item of the framework should support the item it is connected to. The circle connecting the learning content, the gamification components and mechanics, and the ecosystem represents the dependent relationship between the four entities. The framework accounts for problems that could arise during an implementation. For example, if a target behaviour is not being satisfactorily achieved, theory should be considered to determine the problem, and changes should be made to the components and mechanics as necessary.

Figure 6.2 provides an outline of the range of ways the gamified instructional design could be applied to a gamified EFL course. The framework shows example goals, theories, learning content, and the most viable components and mechanics. For an EFL course, the goals will generally be to increase a certain behaviour or target a performance goal; for example, increase EFL ability, increase test scores, increase FL motivation, or increase class participation. For an EFL course, the theory will probably consider how FL acquisition occurs and how specific types of motivation can be targeted. SLA and motivation are complex constructs; therefore, multiple relevant theories should be considered to correctly target specific goals and avoid

negative outcomes. The learning content used in an EFL course would generally be some type of task, skills practice activity, or test. The gamification components and mechanics should be chosen and calibrated to encourage the students to interact with the learning content in a way that is suitable for achieving the goals of the implementation.

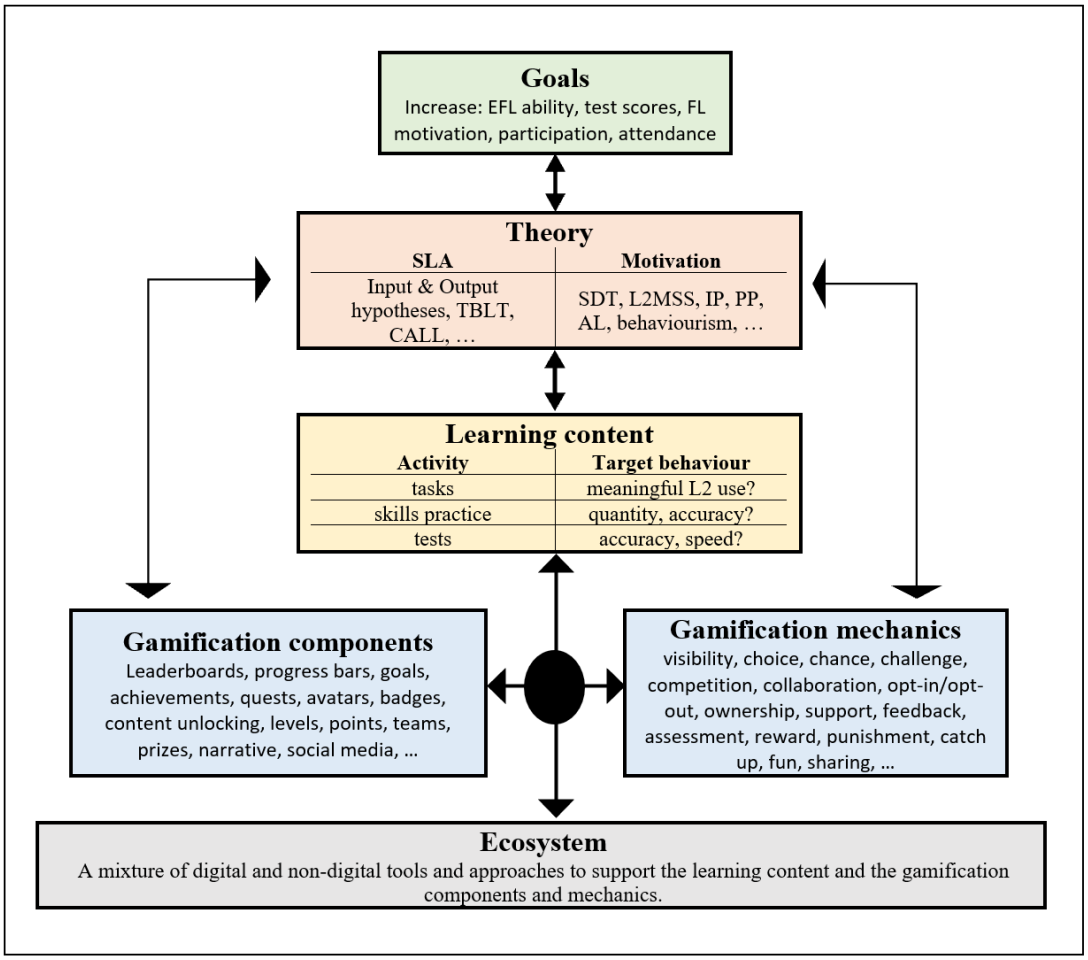


Figure 6.2. The gamified instructional design framework in EFL.
Note. IP = international posture, PP = positive psychology, AL = autonomous learning

Figure 6.2 shows the most viable components and mechanics for gamified instructional design in an EFL context. Each gamification component can incorporate other components and have connections to multiple gamification mechanics. Most of the gamification components and mechanics are borrowed from The Pyramid of Elements (Werbach & Hunter, 2012). However, based on the findings of this study, some new gamification components and mechanics are included. The following

points explain the new components and mechanics, and why they should be included in the gamified instructional design framework. The descriptions of the components and mechanics are purposely broad to allow teachers the freedom to envision and calibrate them in ways suitable for their context.

- Narrative: The Pyramid of Elements (Werbach & Hunter, 2012) classifies narrative as a dynamic. While it is true that it could be considered a dynamic because the participants of a gamified system create their own narrative while using the system, it seems more probable that a specific narrative would be a component that teachers encapsulate their gamified course in.

- Visibility: this refers to how visible the gamification components and mechanics will be to individual students and to the entire community. Depending on the choice of mechanics and components, students may be very aware or completely unaware that their behaviour is being manipulated using gamification. Specific aspects of a component, such as one's leaderboard rank, could be visible to everyone, or only visible to the individual.

- Choice: this refers to the amount of choice students have. For example, can students make choices about the learning content activities, or the members of their team?

- Opt-in/opt-out: can the student opt in and opt out of different aspects of the gamified course, before and during the course. This mechanic could be important to deal with students who feel uncomfortable about the use of a public leaderboard in their class. If they feel uncomfortable, their name could be removed from the leaderboard.

- Ownership: can students take ownership of their work, badges, and prizes, and access them outside of the gamified course? Student motivation to perform well in-class could increase if they feel that they still own their work and achievements, even after class has finished.

- Sharing: can students share their work, badges, and prizes through social media? Student motivation to perform well in-class could increase if they can share their work and achievements, inside and outside of class.

- Support: how much support will the teacher provide to the students during the gamified course? Will the teacher help the students to complete quests that are difficult, or will the students be expected to complete all aspects of the gamified course independent of the teacher?
- Assessment: how will students' completed work be assessed; automatic, peer, or teacher?
- Punishment: how much will undesirable behaviour be punished? The teacher should consider how gamification components such as leaderboards could inadvertently punish low performing students.
- Catch up: is it possible for low performers to catch up?
- Fun: will the components be calibrated to provide a fun environment?

The design of a gamified course should account for the fact that different students have different learning preferences (Tabatabaei & Mashayekhi, 2013) and different students will respond differently to certain components and mechanics (Buckley et al., 2017; Garland, 2015). The bidirectional arrows between the gamification components and mechanics boxes, and the theory box, show how the theory should guide the gamification, and warn that the gamification will impact the theory. Finally, the wide range of CALL-related and gamification supported websites allow the ecosystem to be highly digital. Section 2.3.2 detailed the benefits of using digital technology in gamified EFL instructional design implementations.

A gamified instructional design implementation will probably target multiple goals. Figure 6.3 shows how the gamified instructional design framework could be utilised to target one goal. The goal is to increase the students' TOEIC scores. There are many approaches to achieve the goal, but the implementation in Figure 6.3 aims to achieve the goal through extensive reading that is recommended by Krashen's (2009) reading hypothesis; the literature (i.e., Krashen & Mason, 2015; Nation, 2014) supports the approach. The aim is to increase the amount of words the students read in English. A progress bar that has a goal for the students to achieve is used. Figure 6.3 explains how the mechanics are calibrated to achieve the target behaviour. Goal-

setting theory (Locke & Latham, 2006) supports the components and mechanics and suggest that the students will be motivated to achieve the goal of the progress bar. Behaviourism provides a warning that once the goal has been achieved, the students will stop reading. Therefore, the goals should be created in order to maintain performance.

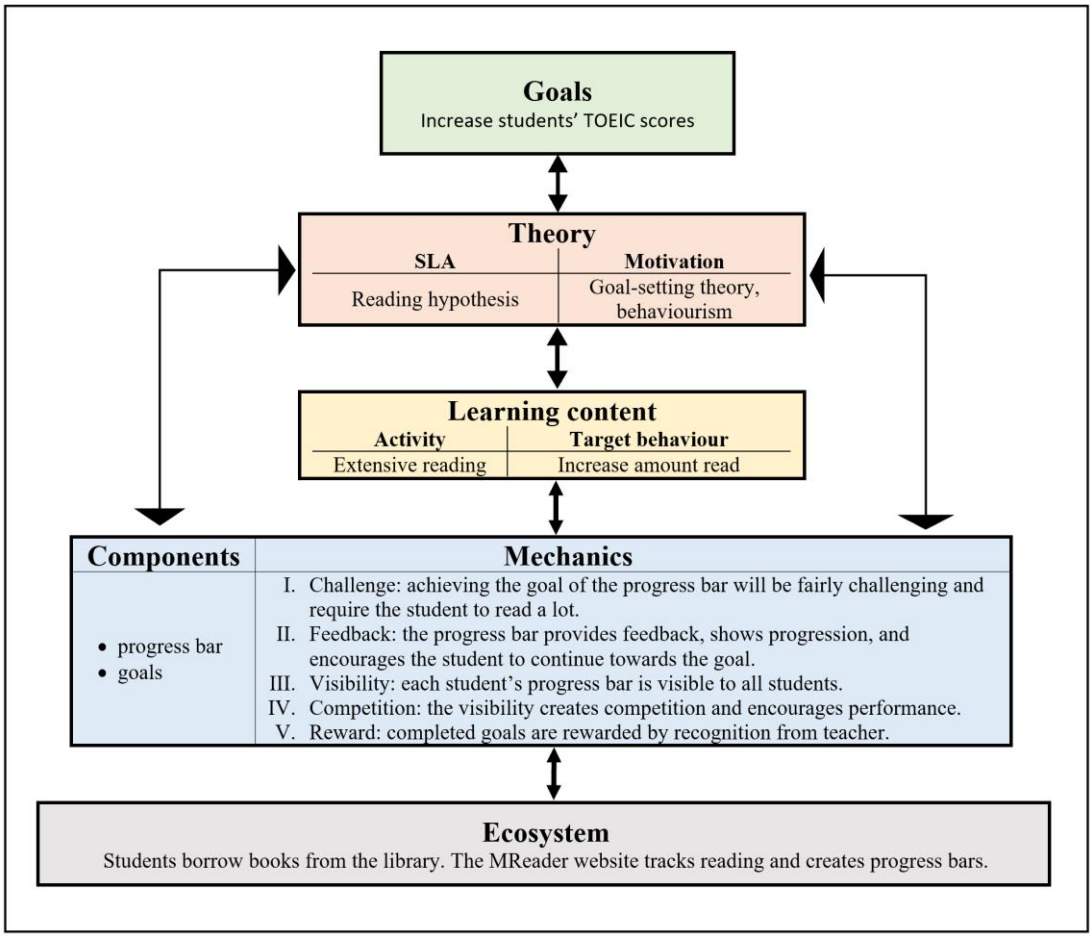


Figure 6.3. The gamified instructional design framework example.

Since Sheldon (2012) first introduced gamified instructional design, researchers have looked into leaderboards and quests, but rarely in a foreign language learning context, and seldom in the Japanese context. The new framework is a significant contribution to the gamified instructional design research field as it is the first framework designed to account for the wide range of linguistic and psychology factors at play in a FL environment. The framework can be used by researchers and teachers to guide their gamified instructional design implementations.

6.2.2 Pedagogical Implications

The findings about QBL are significant as they explore deeper and wider than any other enquiries in QBL. One of the main implications stemming from this study is the suitability of QBL for EFL pedagogy. The results showed that the quests supported FL acquisition and motivation. Applying SDT to the design of the quests was seen to be effective. The three tenets of SDT provided a framework to design quests: (1) foster competence through suitably challenging activities; (2) foster autonomy through quest choice; and (3) foster relatedness through meaningful collaboration. Incorporating positive psychology, the LTMSS framework, and the international posture construct was also seen to have positive effect on intrinsic FL motivation. As QBL generally occurs over at least a semester period, not every quest has to foster deep competence, autonomy, relatedness, and meaningfulness for the QBL system to successfully increase the students' intrinsic FL motivation. However, the design of the QBL should avoid situations in which competence, autonomy, and relatedness are negatively impacted.

The literature (e.g., Burden, 2002; Ebata, 2008; Kobayashi & Brown, 2003; O'Donnell, 2003; Woodrow, 2006) argues that Japanese foreign language learners lack self-confidence in their ability to use foreign languages. The literature (e.g., Benson, 1991; Burden, 2002; Falout, 2004; Kikuchi & Sakai, 2009; Matsuda & Gobel, 2004; Tsuchiya, 2006) attributes the lack of communicative confidence to Japan being a country that lacks opportunities to use English in meaningful ways, and to the education system that uses outdated pedagogical approaches that emphasise test performance over communicative ability. This has also led to low intrinsic motivation to learn EFL. The quests successfully increasing intrinsic FL motivation is significant due to the wide range of positive learning outcomes associated with students who report high levels of self-determined motivation; for example, higher levels of engagement with learning activities (e.g., Ma, 2009; Spratt et al., 2002; Tanaka, 2009), more willingness to communicate (e.g., Nishida, 2012); more frequent use of, and better L2 performance (e.g., Ehrman, 1996; Goldberg & Noels, 2006; Noels et al., 1999, 2001; Vandergrift, 2005), less anxiety, higher positive attitudes towards language learning and elevated feelings of self-efficacy (e.g., Ehrman, 1996; Noels, 2001b; Schmidt, Boraie, & Kassabgy, 1996).

The importance of choice emerged as an important part of QBL. Quest choice and allowing the participants a choice about how they complete their quests is a powerful technique to motivate students. However, some concerns about quest choice arose from the findings of this study. The concerns have implications for teachers looking to utilise QBL. The concerns show that teachers need to balance their desire for students to complete certain quests while minimising any potential negative effects on autonomy.

First, when presented with multiple quests, many of the participants often chose to do quests that looked easy or quick to finish. The participants choosing easy to complete quests over other quests that may be more beneficial raises concerns about the quality of the learning outcomes. However, the participants choosing quests that they can confidently complete could foster a feeling of competence and lead to an increase in intrinsic FL motivation. Second, some of the participants avoided doing the quests that required collaboration because they did not want to work with other people. If a participant avoids a quest that was designed to foster intrinsic FL motivation through relatedness, they are limiting the effectiveness of the QBL. However, there are various reasons why a participant might not want to work with someone else. Forcing a participant to complete collaborative quests when they do not want to could negatively affect their feelings of competence, autonomy, and relatedness. For this study, the findings suggested that, if the participants were provided more time in class to work on the quests with the other participants, the chance of the participants collaborating would have increased.

QBL could be its own stand-alone approach to EFL pedagogy. QBL does not require a leaderboard, but gamification components such as levels and points were shown to be useful. The levels controlled the delivery of the content and the points provided feedback and directed the participants to the activities that the teacher considered important. Teachers looking to implement QBL could follow these steps:

1. Decide the goals of the QBL implementation; the goals will be context specific.

2. Design the ecosystem for quest delivery and submission; the ecosystem will determine what type of quests can be utilised.
3. Design the context specific quests; quest design should be based on content-specific, psychological, and gamification theory.
4. Value the quests; some quests will be more important, difficult, or time consuming than others. Points can be used to show the value of a quest.
5. Sort the quests; quests can be sorted based on their value and/or theme. Quests could be sorted into thematic levels that the students progress through.
6. Allow autonomy; autonomy can be fostered through quest choice and allowing students enough freedom to be creative about how they complete their chosen quests.
7. Integrate feedback; make sure the students share their completed work and receive meaningful feedback.
8. Assess the quests and the QBL system; make changes when necessary.

Leaderboards are a common gamification component, but they are not a necessary gamification component. The results of the study showed that the use of a leaderboard in an educational setting presents a number of paradoxes: (1) they encourage performance but limit performance; (2) increased performance is good for L2 acquisition, but if the performance does not reflect quality learning outcomes it is not good; (3) the points system provides useful feedback, but also leads to a feeling of being controlled; (4) motivating for high rank students but demotivating for low rank students; and (5) even though bad for low rank students, they potentially completed more work than they would have done without a leaderboard. Teachers looking to utilise leaderboards should consider their positive and negative aspects.

The point system encouraging performance but also limiting performance is important for the EFL context due to the way L2 acquisition occurs. SLA theory (e.g., Krashen, 2009; Swain, 1985) encourages large amounts of comprehensive input and output for L2 acquisition to efficiently occur. If a gamification component encourages more usage of the FL, it is beneficial for FL acquisition. If it restricts

performance, it is detrimental to FL acquisition. For this study, the point system represented the minimum amount of work the teacher wanted the participants to complete. However, the participants were encouraged to do the learning activities English Central and extensive reading as much as possible as they are beneficial for their EFL proficiency. Therefore, even though the point system successfully encouraged performance up to a certain point, it also possibly negatively impacted performance. Teachers also need to be careful about what the leaderboard is rewarding. The results of the study show that some of the participants felt that the points reward quantity over quality. The design of the point system should balance the quantity and quality of work.

Traditional university courses allow students to delay completing their homework until just before the deadline. Learning a language does not align to this type of performance as it is a slower process that requires time for the students to process what they have learned. Therefore, the point system was designed to encourage the participants to work on their EFL skills consistently over the semester. The point system provided useful feedback and encouraged the participants to work on their English skills weekly. The point system lacked a mechanism to allow the low-performing students to catch up. However, this could have had a negative impact on the high performing participants if they knew they could have slacked off for a couple of weeks and improve their score in following weeks. The results of the study show the complexity of leaderboard design in an educational setting.

Prior to the study, the researcher assumed that, by applying negative pressure to the low performing participants, the leaderboard would make those students feel compelled to complete all their homework. The results of the study showed that the low-ranked participants did not complete all their homework. However, if there was no leaderboard, maybe the amount of work they completed would have been drastically lower. The leaderboards fostered some negative emotions and attitudes for the low performing participants. As the positive psychology literature (e.g., Seligman & Csikszentmihalyi, 2000) points out the importance of fostering positive emotions rather than negative emotions, a teacher looking to use a leaderboard needs to find the balance between maximum performance and minimum negative emotions. Compared to the low ranked participants, the higher ranked participants were

motivated to maintain their rank, thought the leaderboard was fun, and had stronger positive attitudes towards the learning activities. Recently, more sophisticated leaderboards (e.g., Barata et al., 2017) are being used in which the participants can only see the rankings of the participants who are two places higher than them. This would appear to limit the overwhelming negative feeling of appearing at the bottom of a leaderboard.

6.2.3 Methodological Implications

The two main areas of research enquiry were performance and motivation. Due to the complexity of motivation, it is a harder to accurately measure than performance. This is because the participants are probably not cognisant of the different types of motivation that drive them and because they have varying types of motivation concurrently (Dörnyei & Ushioda, 2011). The complexity of motivation has led researchers to mixed methods research as it allows for more thorough enquiry to occur rather than a simple quantitative or qualitative approach. This study applied a unique mixed methods approach to understand the motivation surrounding educational-based gamification. The approach was unique because of the five data collection instruments. Each instrument provided important insight, combined, the instruments allowed the findings to be triangulated.

The semi-structured interviews were primarily used as part of the mixed methods design to triangulate the findings that emerged from the other data collection tools. The pre- and post-test nature of the LLOS did not capture the ebbs and flows over the 14-week period. However, the periodic leaderboard questionnaire and the quest diaries were able to determine how the participants were responding to the leaderboard and the quests during the middle of the study period. This pragmatic approach proved beneficial as it allowed specific aspects to be closely targeted and it allowed other aspects to freely emerge. The triangulation of results provided support to the reliability and validity of the findings.

Gamification implementations generally focus on motivating a specific behaviour. The performance-related data clearly measured how much work each of the

participants completed. The measures in this study were context specific, and the activities were language learning activities. In other contexts, there will be other measures of performance. The LLOS provided a longitudinal perspective on how motivation changed. The LLOS data clearly aligned the leaderboard to external regulation which was an important finding from this study. Even though the LLOS did not specifically represent competence, autonomy, and relatedness, it did clearly show how the participants' intrinsic FL increased. It was the first time that the LLOS was applied to gamification research. It could be improved if rewritten for the Japanese EFL context.

The leaderboard questionnaire was designed specifically for this study. The leaderboard questionnaire periodically collected data about the participants' emotions and attitudes towards the leaderboards. The leaderboard questionnaire in itself collected mixed methods data. After the participants chose which emotions represented their feelings towards the leaderboards, they explained in their own words why they chose the emotions. Some of the findings were triangulated when the participants responded to the Likert scale statements about their attitudes towards the leaderboards. Exploring emotions using the control-value theory of achievement emotions (Pekrun et al., 2007) was insightful as it showed the relationship between emotions and performance. The attitude section of the leaderboard questionnaire could be improved by also targeting aspects of the point system.

The quest diaries were also designed specifically for this study. They allowed the participants to freely share their opinions of the quests in a non-judgemental setting. Allowing the participants to choose which language they wanted to convey their message was also beneficial for two reasons: first, the data collection served as an extra language learning experience if the participants wanted; and, second, if the students did not know how to express themselves using English or if they preferred, they could use Japanese. The quantitative content analysis successfully allowed for a large amount of qualitative data to be refined to easy to understand themes and provided robust and insightful support for the quantitative LLOS findings.

6.3 Limitations

The findings of the study come with limitations. Gamification research is context-specific. Therefore, replicating the same findings with different groups of participants could be difficult if the gamification components are calibrated in a different way. For example, this study used a limited point leaderboard that showed the rankings of all the participants. If the leaderboard allowed the participants to accumulate as many points as they liked, or if the leaderboard did not show everyone's ranking, the results could be different. The findings about the quests are also specific to the quests used in this study. Different quests could result in different findings.

The are some limitations associated with the leaderboard questionnaire. One of the findings about how emotions activated behaviour determined activated behaviour based on whether the participant received a full score in the following week. However, this activation dichotomy was possibly too simple because, without the leaderboard, a participant may have done no work. The leaderboard possibly activated them to complete 50% or 80% of their work in the following week. Maybe activation should be determined based on whether a participant's score changed, not whether they got a full score or not. As the researcher had experience teaching the participants prior to the study, based on the performance data, it appears that, even though the low performing participants did not always achieve their weekly goals, their performance increased because of the leaderboard.

The attitude section of the leaderboard questionnaire also had a concerning limitation. The limitation was that there were a greater number of positive statements than negative statements. This could have potentially encouraged the researcher to draw more positive conclusions than negative conclusions from the data. Each positive statement should have had a negative counterpart that measured the same attitude. Due to the limitation, the researcher did not value the attitude data as highly as other data.

Four methodological limitations have been identified. The first is about possible questionnaire fatigue for the participants in Class 1. Questionnaire fatigue could

possibly affect the quality of the responses. The second limitation is about the data analyses. The data analyses focused on the average group data. A deeper analysis of the individual data through multiple case studies could lead to new findings. The third limitation is about the sample (N = 46). For the qualitative data, this sample size was sufficient. However, for the quantitative data analyses, a larger sample size would lead to more reliable results. This study used two groups of participants. One more control group which did not include any gamification components could have been useful to determine if the changes in intrinsic motivation were dependent on the QBL. The sample was a group of Japanese students who were the same age and had a similar EFL ability. The findings of the study provide insight into the use of gamification in a Japanese university EFL context, but they cannot be extrapolated to other groups who have a different age or different nationality. The fourth limitation is about the participants' knowledge that they were partaking in a study that was of importance to their teacher. There was a possibility that the participants felt they should work harder than usual because they were involved in a study.

6.4 Future Directions

The continued adoption and refinement of digital systems designed to influence human behaviour suggest that empirical enquiry into gamification is still in its infancy (Mansoor, 2017). The findings of the study provide useful knowledge to the gamification-related research field and suggest future areas of research interest. Leaderboard-related research has many possible areas of inquiry. First, The LLOS data suggested that the leaderboard increased the participants' external regulation at the cost of identified regulation. Studies that attempt to replicate this finding would provide benefit to the gamification literature.

Second, leaderboards can be calibrated in different ways that could have varying effects on the performance and motivation of students (e.g., Jagušt, 2018). For example, rather than having a completely public leaderboard, the leaderboard only shows the participants who are ranked slightly higher. This could encourage performance while mitigating the risk of demotivating low ranked participants. Studies that examine different types of leaderboard calibration would also provide

benefit to gamification literature.

Third, some literature (e.g., Werbach & Hunter, 2012) suggests that, once the novelty of leaderboards wares off, the extrinsic rewards will not be able to sustain engagement. The findings of the study provided light support for the claim. However, many of the participants agreed that they still felt motivated to maintain and gain leaderboard rank at the end of the 14-week period. Therefore, research examining attitudes towards leaderboards over a longer period, such as one year, or research examining different leaderboard calibrations could provide useful insights into education-based gamification.

Fourth, leaderboards by themselves do not increase intrinsic motivation. However, they can provide support to other activities. It seems likely that they could support the growth of intrinsic motivation for high performing students by recognising competence, providing autonomy through feedback, and encouraging relatedness through the communal leaderboard. It also seems likely that they could negatively impact the intrinsic FL motivation of the low performing students through the constant public disclosure of their failed performances. Using the data set, further research that explores the difference between how FL motivation change for the high performing participants and the low performing participants could provide interesting insight. Future research could look deeper than process-product outcomes by focusing on individual learners and individual surges of motivation; Complex dynamic systems theory (Dörnyei, Ibrahim, & Muir, 2015) could provide interesting insight.

References

Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education? It depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, *61*(2), 217-232. doi:10.1007/s11423-013-9289-2

Ackerman, C. (2019). What is the self-determination theory of motivation? Retrieved from https://positivepsychologyprogram.com/self-determination-theory/

Aldemir, T., Celik, B., & Kaplan, G. (2018). A qualitative investigation of student perceptions of game elements in a gamified course. *Computers in Human Behavior*, *78*, 235-254. doi:10.1016/j.chb.2017.10.001

Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior & Human Performance*, *4*(2), 142-175. doi.org/10.1016/0030-5073(69)90004-X

Allport, G. W. (1937). *Personality: A psychological interpretation*. Oxford, England: Holt.

Allwright, D. (1990). Autonomy in language pedagogy: CRILE Working Paper, 6. *University of Lancaster*. Retrieved from http://repository.upi.edu/9592/5/t_bpran0908382bibliograph.pdf

Amano, I. (1996). *Nihon no kyooiku shisutemu: Koozoo to hendoo* [The Japanese education system: Structure and change]. Tokyo: Tokyo University Press.

Ames, C., & Felker, D. W. (1979). An examination of children's attributions and achievement-related evaluations in competitive, cooperative, and individualistic reward structures. *Journal of Educational Psychology*, *71*(4), 413-420. doi:10.1037//0022-0663.71.4.413

Andrade, M. A., & Williams, K. E. (2009). Foreign language learning anxiety in Japanese EFL university classes: Physical, emotional, expressive, and verbal reactions. *Sophia Junior College Faculty Journal, 29*, 1-24.

Antin, J., & Churchill, E. F. (2011). Badges in social media: A social psychological perspective. In *CHI 2011 Gamification Workshop Proceedings* (pp. 1-4). New York, NY: ACM.

Apple, M. T., Da Silva, D., & Fellner, T. (2013). *Language learning motivation in Japan*. New York, NY: Multilingual Matters.

Ashmore, C., & Nitche, M. (2007). The quest in a generated world. Proceedings of the 2007 Digital Games Research Association. *(DiGRA) Conference: Situated Play*, (pp. 503-509). Tokyo, Japan.

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: autonomy enhancing and suppressing teacher behaviours predict students' engagement in schoolwork. *British Journal of Educational Psychology, 72*, 261-278.

Attali, Y., & Arieli-Attali, M. (2015). Gamification in assessment: Do points affect test performance? *Computers & Education*, *83*, 57-63. doi:10.1016/j.compedu.2014.12.012

Au, S. Y. (1988). A critical appraisal of Gardner's socio-psychological theory of second language (L2) learning. *Language Learning, 38*, 75-100.

Aubrey, S. (2014). Development of the L2 motivational self system: English at a university in Japan. *JALT journal*, *36*(2), 153-174.

Aubrey, S., & Nowlan, A. G. P. (2013). Effect of intercultural contact on L2 motivation: a comparative study. In M. T. Fellner, D. Da Silva & T. Fellner

(Eds.). *Language learning motivation in Japan* (pp. 129-151). Bristol: Multilingual Matters.
Auerbach, C. F., & Silverstein, L. B. (2003). *Qualitative data: An introduction to coding and analysis.* New York: New York University Press.
Baber, W. W. (2015). Course level gamification for higher education TEFL in Japan. Retrieved from https://agon.unime.it/files/2016/04/S0307.pdf
Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving participation and learning with gamification. In *Proceedings of the First International Conference on gameful design, research, and applications* (pp. 10-17). New York: ACM.
Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior*, *71(c)*, 550-585.doi:10.1016/j.chb.2016.08.049
Bellotti, F., Berta, R., Gloria, A. D., & Primavera, L. (2009). Enhancing the educational value of video games. *Computers in Entertainment (CIE)*, *7*(2), 1-23.
Benson, P. (2007). Autonomy in language teaching and learning. *Language teaching*, *40*(1), 21-40. doi:10.1017/S0261444806003958
Benson, M. (1991). Attitudes and motivation towards English: A survey of Japanese Freshmen. *RELC Journal*, *22*, 34-48.
Benson, P. (2011). What's new in autonomy. *The Language Teacher*, *35*(4), 15-18.
Berg, L., & Lune, H. (2012). *Qualitative research methods for the social Sciences* (8th ed.) Boston, Pearson. US: Bloomsburg University of Pennsylvania.
Bernaus, M., & Gardner, R. C. (2008). Teacher motivation strategies, student perceptions, student motivation, and English achievement. *Modern Language Journal, 92*(3), 387-401.
Berwick, R., & Ross, S. (1989). Motivation after matriculation: Are Japanese learners of English still alive after exam hell? *JALT Journal 11*(2), 193-210.
Bielik, P. (2012). Integration and adaptation of motivational factors into software systems. In *Personalized Web-Science, Technologies and Engineering: 11th Spring 2012 PeWe Workshop Modra-Piesok, Slovakia April 1, 2012 Proceedings* (pp. 31-32). Retrieved from https://wiki.fiit.stuba.sk/research/seminars/pewe/ontoparty-2011-2012spring/abstracts/31_bielik.pdf
Bogost, I. (2011). Persuasive games: exploitationware. Retrieved from https://www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php.
Borys, M., & Laskowski, M. (2013). Implementing game elements into didactic process: A case study. *Proceedings of the Management, Knowledge and Learning International Conference 2013* (pp. 819-824). Retrieved from https://ideas.repec.org/s/tkp/mklp13.html
Breen, M. (1987). Learner contributions to task design. *Language learning tasks*, *7(*1), 23-46.
Brown, J. D. (2002). English language entrance examinations: *A progress report. In curriculum innovation, testing and evaluation: Proceedings of the 1st annual JALT pan-sig conference* (pp. 94-104). Kyoto Institute of Technology: JALT.
Brown, J. D., & Yamashita, S. (1995). English language entrance exams at Japanese universities: What do we know about them? *JALT Journal*, *17*(1), 7-30.

Brown, H. D. (1990). M & Ms for language classroom? Another look at motivation. In J. Alatis (Ed.), *Georgetown university round table on languages and linguistics* (pp. 383-393). Washington, DC: Georgetown University Press.

Brown, H. D. (2000). *Principles of Language Learning and Teaching* (4th ed.). White Plains, NY: Addison Wesley Longman, Inc.

Buckley, P., & Doyle, E. (2014). Gamification and student motivation. *Interactive Learning Environments, 24*(6), 1162-1175. doi:10.1080/10494820.2014.964263

Buckley, P., Doyle, E., & Doyle, S. (2017). Game on! Students' perceptions of gamified learning. *Journal of Educational Technology & Society, 20*(3), 1-10.

Burden, P. (2002). A cross sectional study of attitudes and manifestations of apathy of university students towards studying English. *The Language Teacher, 26*(3), 3-10.

Burguillo, J. C. (2010). Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & Education, 55*(2), 566-575.

Buunk, A. P., & Gibbons, F. X. (2006). Social comparison orientation: A new perspective on those who do and those who don't compare with others. In S. Guimond (Ed.), *Social comparison and social psychology: Understanding cognition, intergroup, relationship, and culture* (pp. 15-32). New York: Cambridge University Press.

Bursztyn, L., & Jensen, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics, 130*(3), 1329-1367. doi:10.1093/qje/qjv021

Çakıroğlu, Ü., Başıbüyük, B., Güler, M., Atabay, M., & Memiş, B. Y. (2017). Gamifying an ICT course: Influences on engagement and academic performance. *Computers in Human Behavior, 69*(1), 98-107. doi:10.1016/j.chb.2016.12.018

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*(3) doi:10.2307/1170677

Cameron, J. (2001). Negative effects of reward on intrinsic motivation - A limited phenomenon: comment on Deci, Koestner, and Ryan (2001). *Review of Educational Research, 71*(1), 29-42. doi:10.3102/00346543071001029

Carreira, J. M. (2006). Relationships between motivation for learning English and foreign language anxiety: A pilot study. *JALT Hokkaido Journal, 10*, 16-28.

Carreira, J. M. (2012). Motivational orientations and psychological needs in EFL learning among elementary school students in Japan. *System, 40*(2), 191–202. http://dx.doi.org/10.1016/j.system.2012.02.001

Carroll, J. B. (1958). A factor analysis of two foreign language aptitude batteries. *Journal of General Psychology, 59*, 3-19.

Charles, D., Charles, T., McNeill, M., Bustard, D., & Black, M. (2011). Game-based feedback for educational multi-user virtual environments. *British Journal of Educational Technology, 42*(4), 638-654. doi:10.1111/j.1467-8535.2010.01068.x

Charsky, D. (2010). From edutainment to serious games: A change in the use of game characteristics. *Games and Culture, 5*(2), 177-198.

Chen, C. A. (2003). Constructivist approach to teaching: Implications in teaching computer networking. *Information Technology, Learning and Performance Journal, 21*(2), 17-27.

Cheong, C., Cheong, F., & Filippou, J. (2013). 'Quick quiz: A gamified approach for enhancing learning'. In J.-N. Lee, J.-Y. Mao & J. Thong (Eds.) *Pacific Asia conference on information systems (PACIS 2013)* (pp. 1-14). Jeju Island, Korea

Chou, Y. (2015). *Actionable gamification: Beyond points, badges, and leaderboards.* Octalysis Media.

Clark, G. (2010). Saving Japan's universities. Retrieved from https://www.japantimes.co.jp/opinion/2010/08/17/commentary/saving-japans-universities/#.XWuiOSj7R3s

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-449.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*(2), 21-29. https://doi.org/10.1007/BF02299088

Clark, R. E. (2012). *Learning from media. Arguments, analysis, and evidence (*2nd ed). Information Age Publishing.

Clément, R., Gardner, R. C., & Smythe, P. C. (1977). Motivational variables in second language acquisition: A study of Francophones learning English. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement, 9*(2), 123-133.

Clément, R., & Kruidenier, B. (1983). Orientations in second language acquisition: 1. The effects of ethnicity, milieu, and target language on their emergence. *Language Learning, 33*(3), 273–291.

Codish, D., & Ravid, G. (2014). Personality based gamification-educational gamification for extroverts and introverts. In Y. Eshet-Alkalai, A. Caspi, N. Geri, Y. Kalman, V. Silber-Varod, & Y. Yair (Eds.) *Proceedings of the 9th CHAIS conference for the study of innovation and learning technologies: Learning in the technological Era* (pp. 36-44). Retrieved from https://www.openu.ac.il/innovation/chais2014/download/E2-2.pdf

Codish, D., & Ravid, G. (2017). Gender moderation in gamification: Does one size fit all? Retrieved from http://hdl.handle.net/10125/41398

Cofer, C. N., & Petri, H. L. (2001). *Motivation: Behaviour.* Retrieved from https://www.britannica.com/topic/motivation

Cohen, L., Manion, L., & Morrison K. (2011). *Research methods in education* (7th ed.). New York: Routledge.

Comanaru, R., & Noels, K. A. (2009). Self-Determination, motivation, and the learning of Chinese as a heritage language. *Canadian Modern Language Review, 66*(1), 131-158. doi:10.3138/cmlr.66.1.131

Conley, K., & Donaldson, C. (2015). Gamification: The measurement of benefits. In T. Reiners & L. C. Wood (Eds.), *Gamification in education and business* (pp. 673–688). Springer International Publishing.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686. doi:10.1016/j.compedu.2012.03.004

Cook, G. (2018). Top 10 gamification mistakes to avoid. Retrieved from https://www.plecto.com/blog/gamification/top-10-gamification-mistakes-to-avoid/

Coonradt, C. (2012). *The game of work.* Layton, Utah: Gibbs Smith.

Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. New York: Sage.
Cózar-Gutiérrez, R., & Sáez-López, J. M. (2016). Game-based learning and gamification in initial teacher training in the social sciences: an experiment with MinecraftEdu. *International Journal of Educational Technology in Higher Education, 13*(1), 2-17. doi:10.1186/s41239-016-0003-4.
Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: SAGE.
Creswell, J. W. (2015). *A concise introduction to mixed methods research.* Thousand Oaks, California: SAGE.
Crookes, G. & Schmidt, R. (1991). Motivation: Reopening the research agenda. *Language Learning, 41*(4), 469-512.
Csizér, K., & Kormos, J. (2009). Learning experiences, selves and motivated learning behaviour: a comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 98-117). Bristol: Multilingual Mattters.
Curtis, J. R., Wenrich, M. D., Carline, J. D., Shannon, S. E., Ambrozy, D. M., & Ramsey, P. G. (2001). Understanding physicians' skills at providing end-of-life care. *Journal of General Internal Medicine, 16*(1), 41-49. doi:10.1046/j.1525-1497.2001.00333.x
Dam, L. (1995). *From theory to classroom practice*. US: Authentik.
Davis, K. (2007). *Chore wars.* Retrieved from http://www.chorewars.com/
Day R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.
DeCharms, R. (1984). Motivation enhancement in educational settings. In R. Ames & C. Ames (Eds.), *Research on motivation in education: Student motivation* (pp. 275-310). Orlando, FL: Academic Press.
Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*(3-4), 325-346. doi:10.1080/00461520.1991.9653137
Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627-668. doi:10.1037//0033-2909.125.6.627
De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education, 75*(4), 82-91. doi:10.1016/j.compedu.2014.01.012
De Schutter, B., & Abeele, V. V. (2014). *Gradequest - Evaluating the impact of using game design techniques in an undergraduate course*. Retrieved from http://www.fdg2014.org/papers/fdg2014_paper_07.pdf
Deterding, S. (2012). Gamification: Designing for motivation. *Interactions, 19*(4), 14-17. doi:10.1145/2212877.2212883
Deterding, S., Sicart, M., Nacke, L. E., O'Hara, K., & Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. In D. S. Tan, S. Amershi, B. Begole, W. A. Kellogg & M. Tungare (Eds.), *CHI Extended Abstracts* (pp. 2425-2428). Retrieved from http://gamification-

research.org/wp-content/uploads/2011/04/01-Deterding-Sicart-Nacke-OHara-Dixon.pdf

Devers, C. J., & Gurung, R. A. R. (2015). Critical perspective on gamification in education. In T. Reiners & L. C. Wood (Eds.), *Gamification in education and business* (pp. 417–430). London: Springer.

Diamond, L., Tondello, G. F., Marczewski, A., Nacke, L. E., & Tscheligi, M. (2015). The HEXAD gamification user types questionnaire: Background and development process. In *Workshop on personalization in serious and persuasive games and gamified interactions*. London, UK.

Dichev, C., Dicheva, D., Angelova, G., & Agre, G. (2015). From gamification to gameful design and gameful experience in Learning. *Cybernetics and Information Technologies, 14*(4), 80-100. doi:10.1515/cait-2014-0007

Dina, A., & Ciornei, S. (2013). The advantages and disadvantages of computer assisted language learning and teaching for foreign languages. *Procedia - Social and Behavioral Sciences, 76*(1), 248-252. doi:10.1016/j.sbspro.2013.04.107

Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*(1), 380-392. doi:10.1016/j.compedu.2012.12.020

Dörnyei, Z. (1994). Understanding L2 motivation: On with the challenge! *Modern Language Journal, 78*(4), 515-523.

Dornyei, Z. (1990). Conceptualizing motivation in foreign language learning. *Language Learning, 40*(1), 45-78.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition.* Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. and Csizér, K. (2002) Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied Linguistics, 23*(4), 421-462.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Harlow: Longman.

DuBravac, S. (2012). Game mechanics for classroom engagement. In *Increasing student engagement and retention using immersive interfaces: Virtual worlds, gaming, and simulation* (pp. 67-94). Bradford, United Kingdom: Emerald Group Publishing Limited.

Duolingo. (2018). Duolingo for schools. Retrieved from https://www.duolingo.com/

Dweck, C. S. (1999). Self-theories: Their role in motivation, personality, and development (Essays in social psychology). New York, NY: Psychology Press.

Ebata, M. (2008). Motivation factors in language learning. *The Internet TESL Journal, 14*(4), Retrieved from http://iteslj.org/Articles/Ebata-MotivationFactors.html

EF EPI (2018). EF English Proficiency Index. Retrieved from https://www.ef.com/wwen/epi/

Ehrman, M. E. (1996). An exploration of adult language learner motivation, self-efficacy and anxiety. In R. L. Oxford (Ed.), *Language learning motivation: Pathways to the new century* (pp. 81-103). Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center

Ellis, R. (2003). *Task-based language learning and teaching*. UK: Oxford University Press.

Ely, C. (1986). Language learning motivation: A descriptive and causal analysis. *Modern Language Journal, 70*, 28-35.
Eng, D. (2019). Game-Based Learning vs Gamification. Retrieved from https://www.universityxp.com/blog/2019/9/4/game-based-learning-vs-gamification
ETS (2017). Test and score data summary for TOEFL iBT tests. Retrieved from https://www.ets.org/s/toefl/pdf/94227_unlweb.pdf
Evans, M., Jennings, E., & Andreen, M. (2011). Assessment through achievement systems: A framework for educational game design. *International Journal of Game-Based Learning, 1*(3), 16-29.
Falout, J., & Maruyama, M. (2004). A comparative study of proficiency and learner demotivation. *The Language Teacher*, *28*(8), 3-9.
Farber, M. F. (2016). *3 Ways to use game-based learning edutopia.* Retrieved from https://www.edutopia.org/article/3-ways-use-game-based-learning-matthew-farber
Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* East Norwalk, CT: Appleton-Century-Crofts.
Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed). Thousand Oaks, CA: Sage.
Fisher, D., Hafner, C., & Young, J. (2006). Integrating independent learning: lessons learned and implications for the classroom. In D. Gardner (Ed.), *Learner autonomy 10: Integration and support* (pp. 33- 55). Dublin: Authentik.
Flink, C., Boggiano, A. K., & Barrett, M. (1990). Controlling teaching strategies: undermining children's self-determination and performance. *Journal of Personality and Social Psychology, 59*(5), 916-924. doi: 10.1037/0022-3514.59.5.916
Fotaris, P., Mastoras, T. , Leinfellner, R., & Rosunally, Y., (2016) Climbing up the leaderboard: An empirical study of applying gamification techniques to a computer programming class. *Electronic Journal of E-Learning*, *14*(2) 94-110.
Frederick, S., & Loewenstein, G. (1999). Hedonic adaptation. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 302-329). New York, NY: Russell Sage Foundation.
Freud, S. (1966) *The complete introductory lectures on psychoanalysis*. New York: Norton.
Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of economic surveys*, *15*(5), 589-611. doi:10.1111/1467-6419.00150
Fried, C. B. (2008). In-class laptop use and its effects on student learning. *Computers & Education*, *50*(3), 906-914. doi:10.1016/j.compedu.2006.09.006
Fuchs, M., Fizek, S., Ruffino, P., & Schrape, N. (2015). *Rethinking gamification*. Lüneburg: Meson press.
Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95*(1), 148-162.
Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second-language acquisition. *Canadian Journal of psychology revue canadienne de psychologie, 13*(4), 266-272. doi:10.1037/h0083787

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House Publishers.
Gardner, R. C., & Smythe, P. C. (1975). Second language acquisition: A social psychological approach. *Research Bulletin No. 332*. Department of Psychology, University of Western Ontario, London, Canada.
Gardner, R. C. (1979). Social-psychological aspects of second language acquisition. In H. Giles & St. R. Clair (Eds.), *Language and social psychology* (pp. 193-220). Oxford: Blackwell.
Gardner, R.C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
Gardner, R. C. (2010). Second language acquisition: A social psychological perspective. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 204-216). Oxford: Oxford University Press.
Gardner, R. C., & MacIntyre, P. D. (1991). An instrumental motivation in language study: Who says it isn't effective. *Studies in Second Language Acquisition, 13*, 57-72.
Gardner, R.C., & MacIntyre, P.D. (1993). A student's contributions to second language learning. Part II: Affective variables. *Language Teaching 26*(1), 1−11.
Garland, C. M. (2015). *Gamification and implications for second language education: A meta analysis.* Retrieved from https://pdfs.semanticscholar.org/9c92/fa00401e1fe07a02090db482a1f3bb6a3b98.pdf
Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment, 1*(1), 20-20. doi:10.1145/950566.950595
Giles, H., & Byrne, J. L. (1982) An intergroup approach to second language acquisition. *Journal of Multilingual and Multicultural Development, 3*, 17-40. doi:10.1080/01434632.1982.9994069
Glover, I. (2013). Play as you learn: gamification as a technique for motivating learners. In J. Herrington, A. Couros & V. Irvine (Eds.). *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2013* (pp. 1999-2008). Retrieved from http://shura.shu.ac.uk/7172/
Goehle, G. (2013). Gamification and web-based homework. *Primus*, *23*(3), 234-246. doi:10.1080/10511970.2012.736451
Goldberg, E., & Noels, K. A. (2006). Motivation, ethnic identity, and post-secondary education language choices of graduates of intensive French language programs. *The Canadian Modern Language Review*, *62*, 423-447.
González, C., & Area, M. (2013). Breaking the rules: Gamification of learning and educational materials. In *Proceedings of the 2nd international workshop on interaction design in educational environments* (pp. 47-53). doi:10.5220/0004600900470053
Goodwin, B. (2011). Research says… / one-to-one laptop programs are no silver bullet. *Education Leadership, 68*(5) 78-79.
Gorsuch, G. (1998). Yakudoku EFL instruction in two Japanese high school classrooms: An exploratory study. *JALT Journal, 20*(1), 6-32.
Green, C. F. (1999). Categorising motivational drives in second language acquisition. *Language, Culture and Curriculum, 12*(3), 265-279.

Gremmo, M. J., & Riley, P. (1995). Autonomy, self-direction and self access in language teaching and learning: The history of an idea. *System, 23*(2), 151-164.

Growth Engineering (2016). *Countdown to Gamification (History of Gamification Pt.1).* Retrieved from http://www.growthengineering.co.uk/countdown-to-gamification-history-of-gamification-pt-1

Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly, 42*, 55-77. doi: 10.1002/j.1545-7249.2008.tb00207.x

Haaranen, L., Ihantola, P., Hakulinen, L., & Korhonen, A. (2014, March). How (not) to introduce badges to online exercises. In *Proceedings of the 45th ACM technical symposium on computer science education* (pp. 33-38). New York: ACM.

Hakulinen, L., Auvinen, T., & Korhonen, A. (2013). Empirical study on the effect of achievement badges in TRAKLA2 online learning environment. In *2013 learning and teaching in computing and engineering* (pp. 47-54). US: IEEE.

Halan, S., Rossen, B., Cendan, J., & Lok, B. (2010). High score!-motivation strategies for user participation in virtual human development. In *International conference on intelligent virtual agents* (pp. 482-488). Berlin: Springer.

Hamada, Y., & Kito, K. (2007). Demotivation in Japanese high schools. *JALT 2007 Conference Proceedings* (pp. 168-178). Retrieved from http://jalt-publications.org/archive/proceedings/2007/E084.pdf

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? - A literature review of empirical studies on gamification. *2014 Hawaii International Conference on System Sciences* (pp. 3025-3034).

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education, 80*, 152-161. doi:10.1016/j.compedu.2014.08.019

Harwood, T. & Garry, T. (2015). An investigation into gamification as a customer engagement experience environment. *Journal of Services Marketing, 29*(6/7), 533-546.

Hashimoto, Y. (2002). Motivation and willingness to communicate as predictors of reported L2 use: The Japanese ESL context. *Second Language Studies, 20*(2), 29-70.

Haskell, C. C. (2012). Design variables of attraction in quest-based learning (Doctoral dissertation). Retrieved from https://scholarworks.boisestate.edu/td/285/

Haskell, C. C. (2013). *Understanding quest-based learning.* Boise, ID: Boise State University.

Heckhausen, H. (1991). *Motivation and Action.* New York: Springer.

Heider, F. (1958). *The psychology of interpersonal relations*. Hoboken, NJ: John Wiley & Sons Inc.

Heeter, C., Lee, Y. H., Medler, B., & Magerko, B. (2011). Beyond player types: gaming achievement goal. In *ACM SIGGRAPH 2011 Game Papers* (p. 7). New York: ACM.

Herzberg F., Mausner B., & Synderman B. (1959). *The motivation to work*. NY: John Wiley & Sons.

Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. (2016). Engaging Asian students through game mechanics: Findings from two experiment studies. *Computers & Education, 92*(1), 221-236.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological review*, *94*(3), 319-340. doi:10.1037//0033-295x.94.3.319

Higgins, S., Beauchamp, G., & Miller, D. (2007). Reviewing the literature on interactive whiteboards. *Learning, Media and Technology*, *32*(3), 213-225. doi:10.1080/17439880701511040

Hiromori, T. (2003). What enhances language learners' motivation? High school English learners' motivation from the perspective of self-determination theory. *JALT Journal*, *25*(2), 173-186.

Holec, H. (1985). "On autonomy: some elementary concepts". In P. Riley (Ed.), *Discourse and learning* (pp. 173-190). London: Longman.

Hooghiemstra, D. (2016). Gamification in Japan: a critical analysis. Retrived from http://hdl.handle.net/1887/45485

Hoon, H., & Tan, T. M. L. (2008). Organizational citizenship behavior and social loafing: The role of personality, motives, and contextual factors. *The Journal of Psychology, 142*, 89-108. doi:10.3200/jrlp.142.1.89–112

Horwitz, E. K. (1990). Attending to the affective domain in foreign language learning. In S. S. Magnan (Ed.), *Shifting* the *instructional focus* to the *learner* (pp. 15-33). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages

Hunicke, R., Leblanc, M., & Zubek, R. (2004). MDA: A formal approach to game design and game research. *Proceedings of the Challenges in Games AI Workshop, Nineteenth National Conference of Artificial Intelligence* (pp. 1-5). AAAI Press.

Ibanez, M., Di-Serio, A., & Delgado-Kloos, C. (2014). Gamification for engaging computer science students in learning activities: A case study. *IEEE Transactions on Learning Technologies*, *7*(3), 291-301. doi:10.1109/tlt.2014.2329293

Inbar, O., Donitsa-Schmidt, S., & Shohamy, E. (2001). Students' motivation as a function of language learning: The teaching of Arabic in Israel. In Z. Dornyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 297-311). Honolulu: University of Hawaii Press.

Iosup, A., & Epema, D. (2014, March). An experience report on using gamification in technical higher education. In *Proceedings of the 45th ACM technical symposium on Computer science education* (pp. 27-32). New York: ACM.

Izard, C. E. (1991). *The psychology of emotions*. New York: Springer Science & Business Media.

Jia, Y., Liu, Y., Yu, X., & Voida, S. (2017, May). Designing leaderboards for gamification: Perceived differences based on user ranking, application domain, and personality traits. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1949-1960). New York: ACM.

Julkunen, K. (1989). *Situation-and task-specific motivation in foreign-language learning and teaching*. Joensuu: University of Joensuu.

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education.* San Francisco: Pfeiffer.

Kapp, K. M. (2015). *2014 reflections on gamification for learning karl kapp.* Retrieved from http://karlkapp.com/2014-reflections-on-gamification-for-learning/

Kapp, K., Blair, L., & Mesch, R. (2014). *The gamification of learning and instruction fieldbook: Ideas into practice*. San Francisco: John Wiley & Sons.
Kennedy, J. (2018). 3 Economic Challenges Japan Faces in 2019. Retrieved from https://www.investopedia.com/articles/investing/123015/3-economic-challenges-japan-faces-2016.asp
Kikuchi, K., & Sakai, H. (2009). Japanese demotivation to study English: A survey study. *JALT Journal, 31*(2), 183-204.
Kim, B., Park, H., & Baek, Y. (2009). Not just fun, but serious strategies: Using meta-cognitive strategies in game-based learning. *Computers & education, 52*(4), 800-810. doi:10.1016/j.compedu.2008.12.004
Kimura, Y., Nakata, Y., & Okumura, T. (2001). Language learning motivation of EFL learners in Japan: A cross-sectional analysis of various learning milieus. *JALT Journal, 23*(1), 48-67.
Kline, P. (1999). *The handbook of psychological testing* (2nd ed). London: Routledge.
Kochanska, G. (2002). Committed compliance, moral self, and internalization: a mediational model. *Developmental Psychology, 38*(3), 339-351.
Kobayashi, C., & Brown, J. D. (2003). Self-esteem and self-enhancement in Japan and America. *Journal of Cross-cultural Psychology, 34*(5), 567-580.
Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes.* US: Houghton Mifflin Harcourt.
Koga, T. (2010). Dynamicity of motivation, anxiety and cooperativeness in a semester course. *System, 38*(2), 172-184.
Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior, 35*(1), 179-188. doi:10.1016/j.chb.2014.03.007
Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management, 45*, 191-210.
Koster, R. (2005). *A Theory of fun.* Scottsdale: Scottsdale, AZ.
Kozma, R. B. (1994). Will media influence learning: Reframing the debate. *Educational Technology Research and Development, 42*(2), 7-19. https://doi.org/10.1007/BF02299087
Kozma, R. (2000). Reflections on the state of educational technology research and development. *Educational Technology, Research and Development*, 48(1), 5-15. https://doi.org/10.1007/BF02313481
Krashen, S. (1982). *Principles and practice in second language acquisition.* London: Pergamon. doi:10.2307/3586656
Krashen, S. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Portsmouth, NH: Heinemann
Krashen, S. (2009). *Principles and practice in second language acquisition*. Retrieved from http://www.sdkrashen.com/content/books/principles_and_practice.pdf
Krashen, S., & Mason, B. (2015). Can second language acquirers reach high levels of proficiency through self-selected reading? An attempt to confirm nation's (2014) results. *The International Journal of Foreign Language Teaching, 10*(2), 10-19.
Krause, M., Mogalle, M., Pohl, H., & Williams, J. J. (2015). A playful game changer: Fostering student retention in online education with social

gamification. In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale* (pp. 95-102). New York: ACM.

Kubota, R. (1998). Ideologies of English in Japan. *World Englishes, 17*(3), 295-306.

Lambert, W. E. (1955). Measurement of linguistic dominance of bilinguals. The *Journal of Abnormal and Social Psychology, 50*(2), 197-200. doi:10.1037/h0042120

Lambert, W. E. (1956a). Developmental aspects of second-language acquisition: I. Associational fluency, stimulus provocativeness, and word-order influence. *The Journal of Social Psychology, 43*, 83-89.

Lambert, W. E. (1956b). Developmental aspects of second-language acquisition: II Associational Stereotypy, associational form, vocabulary commonness, and pronunciation. *The Journal of Social Psychology, 43*, 91-98.

Lambert, W. E. (1956c). Developmental aspects of second-language acquisition: III. A description of developmental changes. *The Journal of Social Psychology, 43*, 99-104.

Lambert, J., & Ennis, J. (2014). Quest-based learning: A new approach to preservice teacher technology instruction. *In society for information technology & teacher education international conference* (pp. 2895-2900). US: Association for the Advancement of Computing in Education.

Lambert, J., Gong, Y. & Harrison, R. (2015). A study of motivation in a quest-based learning environment. In D. Rutledge & D. Slykhuis (Eds.), *Proceedings of SITE 2015 - Society for information technology & teacher education international conference* (pp. 127-132). Las Vegas, NV, United States: Association for the Advancement of Computing in Education.

Lambert, J. (2017). Differences in student motivation and orientation between quest-based and traditional college-level courses. In P. Resta & S. Smith (Eds.), *Proceedings of society for information technology & teacher education international conference* (pp. 1685-1688). Austin, TX, United States: Association for the Advancement of Computing in Education.

Lam, S., Yim, P., Law, J. S., & Cheung, R. W. (2004). The effects of competition on achievement motivation in Chinese classrooms. *British Journal of Educational Psychology*, *74*(2), 281-296. doi:10.1348/000709904773839888

Landers, R. N., Bauer, K. N., Callan, R. C., & Armstrong, M. B. (2015). Psychological theory and the gamification of learning. In T. Reiners & L. C. Wood (Eds.), *Gamification in Education and Business* (pp. 165-186). Springer International Publishing.

Landers, R. N., & Landers, A. K. (2014). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & gaming*, *45*(6), 769-785.

Leadership-central.com (2018). Motivation theories. Retrieved from https://www.leadership-central.com/motivation-theories.html

Lee, J. J. & Hammer, J. (2011). Gamification in education: What, how, why bother? *Academic Exchange Quarterly*, *15*(2), 146.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of personality and social psychology*, *28*(1), 129-137. doi:10.1037/h0035519

Linnet, J., Møller, A., Peterson, E., Gjedde, A., & Doudet, D. (2011). Dopamine release in ventral striatum during Iowa gambling task performance is

associated with increased excitement levels in gambling disorder. *Addiction 106*, 383–390. doi: 10.1111/j.1360-0443.2010.03126.x
Lister, M. (2015). Gamification: The effect on student motivation and performance at the post secondary level. *Issues and Trends in Educational Technology, 3*(2), 227-232.
Liu, D. (2015). A critical review of Krashen's input hypothesis. *Journal of Education and Human Development*, *4*(4), 139-146. doi:10.15640/jehd
Little, D. G. (1991). *Learner autonomy I: Definitions, issues and problems*. Dublin: Authentik.
Locke, E.A., & Latham, G.P. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice Hall.
Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, *15*(5), 265-268. doi:10.1111/j.1467-8721.2006.00449.x
LoCastro, V. (2001). Individual differences in second language acquisition: Attitudes, learner subjectivity, and L2 pragmatic norms. *System, 29*(1), 69-89.
Lopez, S. (2011). Steve lopez: Disneyland workers answer to 'electronic whip'. Retrieved from http://articles.latimes.com/2011/oct/19/local/la-me-1019-lopez-disney-20111018
Lounis, S., Pramatari, K., & Theotokis, A. (2014). Gamification is all about fun: The role of incentive type and community collaboration. In *Proceedings of ECIS 2014*: Tel Aviv.
Lukmani, Y. M. (1972). Motivation to learn and language proficiency. *Language Learning*, *22*, 261-273.
Lynch, B. K. (2003). *Language assessment and programme evaluation.* Edinburgh: Edinburgh University Press.
Lyubomirsky, S. (2007). *The how of happiness: A scientific approach to getting the life you want.* New York, NY, US: Penguin Press.
Ma, J. H. (2009). Autonomy, competence, and relatedness in L2 learners' task motivation: A self-determination theory perspective (unpublished doctoral dissertation). Honolulu, HI: University of Hawaii at Manoa.
Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design.* Mahwah, NJ: Lawrence Erlbaum.
Malone, T. W. (1980). What makes things fun to learn? Heuristics for designing instructional computer games. In *Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems* (pp. 162-169). New York: ACM.
Mansoor, H. (2017). How social media influences customer behavior to drive traffic and sales. Retrieved from: http://customerthink.com/how-social-media-influencescustomer-behavior-to-drive-traffic-and-sales/
Marczewski, A. C. (2015). *Even ninja monkeys like to play: Gamification, game thinking and motivational design*. US: CreateSpace Independent Publishing Platform.
Marklein, M. B. (2010). Can college students learn as well on iPads, e-books? Retrieved from http://www.usatoday.com/news/education/2010-08-10-ebooklearning10_CV_N.htm.
Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954-969. doi:10.1037//0003-066x.41.9.954

Marshall, E., & Marquier, B. (n.d.). Friedman test in SPSS. Retrieved from https://www.sheffield.ac.uk/polopoly_fs/1.714575!/file/stcp-marshall-FriedmanS.pdf

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370-396.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System, 32*, 21-36.

McCombs, B. L. (1994). Strategies for assessing and enhancing motivation: Keys to promoting self-regulated learning and performance. In H. R. O'Neil & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 49-69). Hillsdale, NJ: Lawrence Erlbaum.

McDaniel, R., Lindgren, R., & Friskics, J. (2012). Using badges for shaping interactions in online learning environments. In *2012 IEEE International Professional Communication Conference* (pp. 1-4). Retrieved from https://ieeexplore.ieee.org/document/6408619

McEown, M. S., Noels, K. A., & Saumure, K. D. (2014). Students' self-determined and integrative orientations and teachers' motivational support in a Japanese as a foreign language context. *System, 45*, 227-241. doi:10.1016/j.system.2014.06.001

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. New York: Penguin Press.

McVeigh, B.J. (2001). Higher education, apathy and post-meritocracy. *The Language Teacher, 25*(10), 29-32.

McVeigh, B. J. (2004). Foreign language instruction in Japanese higher education: The humanistic vision or nationalist utilitarianism? *Arts and Humanities in Higher Education, 3*(2), 211-227. doi:10.1177/1474022204042687

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013). Do points, levels and leaderboards harm intrinsic motivation?: An empirical analysis of common gamification elements. In *Proceedings of the First International Conference on gameful design, research, and applications* (pp. 66-73). New York: ACM.

Mertens, D. M. (2009). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (3rd ed.). Thousand Oaks, CA: Sage.

Meyer, K. (2008). Do rewards shape online discussions. *Journal of Interactive Online Learning, 7*(2), 126-138.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Mollick, E. R., & Rothbard, N. (2014). Mandatory fun: Consent, gamification and the impact of games at work. *The Wharton School research paper series*. New York: The Wharton School Research Paper Series.

Morales, M., Amado-Salvatierra, H. R., Hernández, R., Pirker, J., & Gütl, C. (2016). A practical experience on the use of gamification in MOOC courses as a strategy to increase motivation. In *International Workshop on Learning Technology for Education Challenges* (pp. 139-149). New York: Springer, Cham.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained. *Journal of Mixed Methods Research, 1*(1), 48-76. doi:10.1177/2345678906292462

Muntean, C. I. (2011). Raising engagement in e-learning through gamification. In *Proceedings of the 6th International Conference on Virtual Learning* (pp. 323–329), Cluj-Napoca, Romania, Europe.
Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin, 83*(4), 602-627. doi:10.1037//0033-2909.83.4.602
Myerson, R. B. (1999). Nash equilibrium and the history of economic theory. *Journal of Economic Literature, 37*(3), 1067-1082.
Nagano, Y. (2009). Japans 'exam hell' now reaches into preschool. Retrieved from http://www.csmonitor.com/World/Asia-Pacific/2009/0423/p06s01-woap.html
Nakata, Y. (2006). *Motivation and experience in foreign language learning.* Oxford, UK: Peter Lang.
Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*(2), 1-16.
Nishida, R. (2012). An empirical study on L2 ideal self, international posture, intrinsic motivation, willingness to communicate among EFL Japanese university students. Paper presented at the JACET convention 2012. Aichi, Japan.
Nicholson, S. (2013). Exploring Gamification Techniques for Classroom Management. *Paper: Presented at Games+Learning+Society 9.0, Madison, WI.* Retrieved from http://scottnicholson.com/pubs/gamificationtechniquesclassroom.pdf.
Nicholson, S. (2015). A recipe for meaningful gamification. In T. Reiners & L. C. Wood (Eds.), *Gamification in education and business* (pp. 1-20). Springer International Publishing.
Noels, K. A. (2001a). Learning Spanish as a second language: Learners' orientations and perceptions of teachers' communicative style. Language Learning, *51*, 107–144.
Noels, K. A. (2001b). New orientations in language learning motivation: Towards a model of intrinsic, extrinsic and integrative orientations. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp.43–68). Honolulu, HI: University of Hawai'i, Second language Teaching and Curriculum Center.
Noels, K. A. (2005). Orientations to learning German: Heritage language background and motivational processes. *Canadian Modern Language Review, 62*(2), 285-312.
Noels, K. A. (2013). Promoting the motivation of Japanese learners of English through autonomy, competence, and relatedness. *Language learning motivation in Japan*, *17*(1), 15-34.
Noels, K. A., Clément, R., & Pelletier, L. G. (1999). Perceptions of teacher communicative style and students' intrinsic and extrinsic motivation. *Modern Language Journal, 83*(1), 23-34.
Noels, K. A. Clément, R., & Pelletier, L. G. (2001). Intrinsic, extrinsic and integrative orientations of French Canadian learners of English. *Canadian Modern Language Review, 57*(3), 424-442.
Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why Are you learning a second language? Motivational orientations and self-determination theory. *Language learning*, *50*(1), 57-85. doi:10.1111/0023-8333.00111

Nunan, D. (1997). Designing and adapting materials that encourage learner autonomy. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 192- 203). London: Longman.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. UK: Cambridge University Press.

Ockert, D. (2011). A multivariate analysis of Japanese university student motivation and pedagogical activity preferences. *The Language Teacher, 35*(2), 9-17.

O'Donnell, K. (2003). Uncovering first year students' language learning experiences, their attitudes, and motivations in a context of change at the tertiary level of education. *JALT Journal, 25*(1), 31-62.

O'Donovan, S., Gain, J., & Marais, P. (2013). A case study in the gamification of a university-level games development course. In *Proceedings of the South AfricanInstitute for Computer Scientists and Information Technologists Conference* (pp. 242 251). Retrieved from https://people.cs.uct.ac.za/~jgain/publications/gamification.pdf

Oller, J. W., Jr., Hudson, A. J., & Liu, P. F. (1977). Attitudes and attained proficiency in ESL: A sociolinguistic study of native speakers of Chinese in the United States. *Language Learning, 27*(1), 1-23.

Orosz, G., Farkas, D., & Roland-Levy, C. P. (2013). Are competition and extrinsic motivation reliable predictors of academic cheating? *Frontiers in Psychology*, *4*(1), 87.

Otoshi, J., & Heffernan, N. (2011). An analysis of a hypothesized model of EFL students' motivation based on self-determination theory. *The Asian EFL Journal Quarterly, 13*, 66-86.

Oxford, R. L., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *Modern Language Journal, 78*, 12-28.

Oxford, R. L., & Shearin, J. (1996). Language learning motivation in a new key. In R. L. Oxford (Ed.). *Language learning motivation: Pathways to the new century* (pp. 121-144). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center.

Pae, T. (2008). Second language orientation and self-determination theory. *Journal of language and social psychology*, *27*(1), 5-27. doi:10.1177/0261927x07309509

Pavlov, I. P., & Thompson, W. H. (1902). *The work of the digestive glands.* London : Griffin.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice: the definitive text of qualitative inquiry frameworks and options*. Thousand Oaks, CA: Sage Publications

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz & R. P. Perry (Eds.) *Emotion in education* (pp. 13-36). Dublin: Academic Press.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary educational psychology*, *36*(1), 36-48. doi:10.1016/j.cedpsych.2010.10.002

Pemberton, R., Toogood, S., & Barfield, A. (Eds.). (2009). *Maintaining control: Autonomy and language learning* (Vol. 1). Hong Kong: Hong Kong University Press.

Perry, B. (2015). Gamifying French language learning: A case study examining a quest-based, augmented reality mobile learning-tool. *Procedia - social and behavioral sciences*, *174*(1), 2308-2315. doi:10.1016/j.sbspro.2015.01.892

Philpott, A. (2015a). *Students' reactions to the use of leaderboards in an EFL Class: A gamification-related pilot study. Annual Research Report of the Language Center, 17, 83-99.* Nishinomiya: Japan: Kwansei Gakuin University.

Philpott, A. (2015b). *Developing student autonomy using semi-autonomous self-access learning tasks in EFL classrooms. Kwansei Gakuin University Humanities Review, 20, 65-80.* Nishinomiya, Japan: Kwansei Gakuin University.

Poondej, C., & Lerdpornkulrat, T. (2016). The development of gamified learning activities to increase student engagement in learning. *Australian Educational Computing, 31*(2), 1-16.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment*, *1*(1), 21. doi:10.1145/950566.950596

Prensky, M. (2001). *Digital Game-Based Learning*. New York: McGraw-Hill Psychologist (n.d.). Pavlov's Dogs and Classical Conditioning. Retrieved from https://www.psychologistworld.com/behavior/pavlov-dogs-classical-conditioning

Ramage, K. (1990). Motivational factors and persistence in foreign language study. *Language Learning, 40*, 189-219.

Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. Motivation and Emotion 13(2), 83 -103.

Reeve, J. (2002). Self-determination theory applied to educational settings. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 183-203). Rochester, NY: University of Rochester Press.

Reeve, J., & Deci, E. L. (1996). Elements of the competitive situation that affect intrinsic motivation. *Personality and social psychology bulletin*, *22*(1), 24-33. doi:10.1177/0146167296221003

Reinhardt, J. (2017). Digital gaming in L2 teaching and learning. In C. Chapelle & S. Sauro (Eds.), *Handbook of technology in second language teaching and learning* (pp. 202-216). Hoboken, NJ: Wiley-Blackwell.

Reinhardt, J. (2019). *Gameful second and foreign language teaching and learning: Theory, research, and practice*. Basingstoke, UK: Palgrave Macmillan.

Richter, G., Raban, D. R., & Rafaeli, S. (2015). Studying gamification: the effect of rewards and incentives on motivation. In T. Reiners & L. C. Wood (Eds.), *Gamification in education and business* (pp. 21–46). Springer International Publishing.

Robertson, M. (2010). Can't play, won't play. Retrieved from http://hideandseek.net/2010/10/06/cant-play-wont-play/

Roeckelein, J.E. (2006). *Elsevier's Dictionary of Psychological Theories*. London: Elsevier Science.

Rohlen, T. P. (1983). *Japan's high schools*. Berkeley: University of California Press.

Rubenfeld, S., Sinclair, L., & Clément, R. (2007). Second language learning and acculturation: The role of motivation and goal content congruence. *Canadian Journal of Applied Linguistics, 10*(3), 309-323.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54-67. doi:10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 171-195). New York, NY, US: Routledge/Taylor & Francis Group.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*(4), 344-360. doi:10.1007/s11031-006-9051-8

Ryan, S. (2009). Ambivalence and commitment, liberation and challenge: investigating the attitudes of young Japanese people towards the learning of English. *Journal of Multilingual and Multicultural Development, 30*(5), 405-420.

Saldaña, J. (2015). *The coding manual for qualitative researchers*. New York: Sage.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Basingstoke: Palgrave Macmillan.

Schmidt, R., Boraie, D., & Kassabgy, O. (1996). Foreign language motivation: Internal structure and external connections. In R. Oxford (Ed.), *Language learning motivation: Pathways to the new century* (pp. 9-70). Honolulu, HI: University of Hawai'i.

Schumann, J. H. (1978). The acculturation model for second language acquisition. In R. C. Gingras (Ed.), *Second language acquisition and foreign language teaching* (27-50). Washington: Center for Applied Linguistics

Seligman, M. E. P. (1998). *Learned optimism*. New York: Pocket Books.

Seligman, M. (2008). *The new era of positive psychology*. Retrieved from https://www.youtube.com/watch?v=9FBxfd7DL3E

Seligman, M. E., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5-14. doi:10.1037//0003-066x.55.1.5

Sergis, S., Sampson, D. G., & Pelliccione, L. (2018). Investigating the impact of flipped classroom on students' learning experiences: A self-determination theory approach. *Computers in human behavior, 78*(1), 368-378. doi:10.1016/j.chb.2017.08.011

Shehadeh, A., & Coombe, C. A. (2012). *Task-based language teaching in foreign language contexts: Research and implementation.* Amsterdam: John Benjamins Publishing Company.

Sheldon, L. (2011). *The multiplayer classroom: Designing coursework as a game*. Boston, MA: Cengage Learning.

Sierra, K. (2011). *Pixie dust & the mountain of mediocrity.* Retrieved from https://www.gapingvoid.com/blog/2011/06/07/pixie-dust-the-mountain-of-mediocrity/

Skehan, P. (1989). *Individual differences in second-language learning*. London: Edward Arnold.

Skehan, P. (1998). A cognitive approach to language learning. Oxford: Oxford University Press.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.

Skinner, B. F. (1948). Superstition' in the pigeon. *Journal of Experimental Psychology, 38*(2), 168-172.

Soares Palmer, D. (2010). *Second language pragmatic socialization in world of warcraft*. Retrieved from https://eric.ed.gov/?id=ED520797.

Son, J.-B. (2018). *Teacher development in technology-enhanced language teaching*. London: Palgrave Macmillan.

Son, J.-B., & Windeatt, S. (2017). Teacher training in computer-assisted language learning: Voices of teacher educators. In J.-B. Son & S. Windeatt (Eds.), *Language teacher education and technology: Approaches and practices* (pp. 1-17). London: Bloomsbury Academic.

Spratt, M., Humphreys, G., & Chan, V. (2002). Autonomy and motivation: which comes first? *Language Teaching Research, 6*(3), 245-266.

Stapleton, P. (1995). The role of Confucianism in Japanese education. *Language Teacher: Kyoto JALT, 19*, 13-17.

Sugimoto, Y. (2010). *An introduction to Japanese society*. Cambridge: Cambridge University Press.

Sugita, M., & Takeuchi, O. (2010). What can teachers do to motivate their students? A classroom research on motivational strategy use in the Japanese EFL contest. *Innovation in Language Learning and Teaching, 4*, 21-35.

Sugita McEown, M., & Takeuchi, O. (2014). Motivational strategies in EFL classrooms: How do teachers impact students' motivation? *Innovation in Language Learning and Teaching, 7*, 1-19.

Sullivan, A., Mateas, M., Wardrip-Fruin, N. (2009). Questbrowser: Making quests playable with computer-assisted quest design. In *Proceedings of the 8th Digital Art and Culture Conference*. Retrieved from https://pdfs.semanticscholar.org/467e/220cec7501d2822111565f9040e4313bd95c.pdf

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-256). Rowley, MA: Newbury House.

Swain, M. (2000). The output hypothesis and beyond mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.

Tabatabaei, O., & Mashayekhi, S. (2013). The relationship between EFL learners' learning styles and their L2 achievement. *Procedia-Social and Behavioral Sciences, 70*, 245 253.

Tachibana, Y., Matsukawa, R., & Zhong, Q. X. (1996). Attitudes and motivation for learning English: A cross-national comparison of Japanese and Chinese high school students. *Psychological Reports, 79*, 691-700.

Taguchi, N. (2005). The communicative approach in Japanese secondary schools: Teachers' perceptions and practice. *The Language Teacher, 29*(3), 3-9.

Taguchi, T. (2013). Motivation, attitudes and selves in the Japanese context: A mixed methods approach. In M. Apple, D. Da Silva & T. Fellner (Eds.) *Language learning motivation in Japan* (pp. 169-188). Bristol: Multilingual Matters.

Tan, M., & Hew, K. F. (2016). Incorporating meaningful gamification in a blended learning research methods class: Examining student learning, engagement, and affective outcomes. *Australasian Journal of Educational Technology, 32*(5), 19-34.

Tanaka, H. (2009). Enhancing intrinsic motivation at three levels: the effects of motivational strategies. *JALT Journal, 31*, 227-250.

Tang, S., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology, 9*(5), 365-404. doi:10.1002/acp.2350090502

Tashakkori, A., & Teddlie. C., (2010). *SAGE handbook of mixed methods in social and behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.

Terrill, B. (2008). My coverage of lobby of the social gaming summit. Retrieved from http://www.bretterrill.com/2008/06/my-coverage-of-lobby-of-social-gaming.html

Thiebes, S., Lins, S., & Basten, D. (2014). Gamifying information systems - a synthesis of gamification mechanics and dynamics. In: ECIS, Tel Aviv, Israel. Retrieved from https://aisel.aisnet.org/ecis2014/proceedings/track01/4/

Thulin, L. (2018). The roots of the cambridge analytica scandal: How farmville helped users become comfortable giving away their Facebook data. Retrieved from https://slate.com/technology/2018/03/farmville-helped-sow-the-seeds-of-the-cambridge-analytica-scandal.html.

Tran, A., & Zeckhauser, R. (2012). Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics, 96*(9-10), 645-650. doi: 10.1016/j.pubeco.2012.05.004

Tremblay, P.F., & Gardner, R.C. (1995). Expanding the motivation construct in language learning. *Modern Language Journal, 79*(4), 505–20.

Tsuchiya, M. (2006). Factors in demotivation of lower proficiency English learners at college. *The Kyushu Academic Society of English Language Education (KASELE), 34*, 87-96.

Ushioda, E. (2013). Foreign language motivation research in Japan: An 'insider' perspective from outside Japan. In: M. T. Apple, D. Da Silva & T. Fellner (Eds.) *Language Learning Motivation in Japan* (pp. 1-14). Bristol: Multilingual Matters

Vallerand, R. J. (1983). Effect of differential amounts of positive verbal feedback on the intrinsic motivation of male hockey players. *Journal of Sport Psychology, 5*, 100-107.

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 271-360). New York: Academic Press.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and psychological measurement, 52*(4), 1003-1017. doi:10.1177/0013164492052004025

Vallerand, R. J. and Ratelle, C. F., (2002). Intrinsic and extrinsic motivation: A hierarchical model. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp 128-127). Rochester, NY: The University of Rochester Press.

Van Avermaet, P., & Gysen, S. (2006). From needs to tasks: Language learning needs in a task based approach. *Task-based language education: from theory to practice*, *17*-16. UK: Cambridge University Press.

Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied linguistics, 26*(1), 70-89. doi:10.1093/applin/amh039

Vansteenkiste, M., Aelterman, N., De Muynck, G. J., Haerens, L., Patall, E., & Reeve, J. (2018). Fostering personal meaning and self-relevance: A self-determination theory perspective on internalization. *The Journal of Experimental Education, 86*(1), 30-49. doi: 10.1080/00220973.2017.1381067

Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology, 101*(3), 671-688.
Vroom, V.H. (1964). *Work and motivation*. Oxford, England: Wiley.
Warden, C., & Lin, H. J. (2000). Existence of integrative motivation in Asian EFL setting. *Foreign Language Annals, 33*(5), 535-547. doi:10.1111/j.1944-9720.2000.tb01997.x
Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3-20). Tokyo: Logos International.
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological review, 20*(2), 158-177. doi:10.1037/h0074428
Weinschenk, S. (2013). Use unpredictable rewards to keep behaviour going. Retrieved from https://www.psychologytoday.com/blog/brain-wise/201311/use-unpredictable-rewards-keep-behavior-going
Wells, B. M., & Skowronski, J. J. (2012). Evidence of choking under pressure on the PGA tour. *Basic and Applied Social Psychology, 34*(2), 175-182.
Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business*. Philadelphia, PA: Wharton Digital Press.
Willis, J., & Willis, D. (1996). *Challenge and change in language teaching*. Macmillan: Education Australia.
Willis, D., & Willis, J. (2007). *Doing Task-Based Teaching*. Oxford: Oxford University Press.
Winoto, P., & Tang, T. Y. (2015). From market place to collusion detection: Case studies of gamification in education. In T. Reiners & L. C. Wood (Eds.), *Gamification in education and business* (pp. 277-290). Springer International Publishing.
Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal*. *37*(3). 308-28.
Wu, X. (2003). Intrinsic motivation and young language learners: the impact of the classroom environment. *System, 31*(4), 501-517. doi:10.1016/j.system.2003.04.001
Xu, Y. (2012). Literature review on web application gamification and analytics. Retrieved from http://csdl.ics.hawaii.edu/techreports/11-05/11-05.pdf
Yamashiro, A., & McLaughlin, J. (2001). Relationships among attitudes, motivation, anxiety and English language proficiency in Japanese college students. In P. Robinson, M. Sawyer & S. Ross (Eds.), *Second language acquisition research in Japan* (pp 113- 127). Japan: Japan Association for Language Teaching.
Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, *86*(1), 54-66. doi:10.1111/1540-4781.00136
Yashima, T. (2009). International posture and the ideal L2 self in the Japanese EFL context. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self (*pp. 144-163). Bristol: Multilingual Matters.
Zeidner, M. (1998). *Test anxiety: The state of the art.* New York: Plenum
Zhou, M., Ma, W. J., & Deci, E. L. (2009). The importance of autonomy for rural Chinese children's motivation for learning. *Learning and Individual Differences, 19*(4), 492-498.

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. Sebastopol, CA: O'Reilly Media.
Zimbardo, P. G., Pilkonis, P. A., & Marnell, M. E. (1977). Shyness: What it is, what to do about it. Reading, Mass: Addison-Wesley.

Appendix A: Quest Descriptions

Below is an edited version of all the quests used in this study. All pictures, videos, formatting, further instructions, and examples have been removed.

Level 1: Let's get started

Hello and welcome. This semester is designed like a game. Please read the rules of the game. Your first challenge is to make a blog on Blogger, and then post the URL of your blog to the appropriate thread on our Facebook group. In your first blog post, introduce yourself and add some pictures or photos to make your post more interesting. Also, don't forget to do your extensive reading and English Central homework activities. Your progress will be checked at the beginning of class on Friday

Level 2: Getting to know you

Welcome to Level 2! This is where the fun will begin. There are 3 quests designed to make you think about your life and your goals for this semester. On the left side of the screen you can see the 3 quests. When you have completed 2 of the quests, you may proceed to Level 3 by clicking the 'Change Level' button. You can choose what quests to do. Don't be afraid to try new things. Enjoy the challenge. Be creative. Ask for help if you are not sure about something.

Quest 1: Narrated Presentation about Myself

- Goal: Make a narrated presentation about yourself and your language learning history.
- What is a narrated presentation? It is a digital presentation in which you have recorded your voice and set the timings for each slide so that the presentation can run automatically.
- Length: As long as you like.
- Instructions: Create a narrated presentation, convert it to a video, upload it to YouTube, post it to your blog after your weekly diary entry.

Quest 2: Write: Self-Assessment

- Goal: Write a blog post discussing your strengths and weaknesses. How can you overcome your weaknesses?
- Length: At least 150 words
- Instructions: Post your writing to your blog with a diary entry.

Quest 3: Video: Goals for the semester

- Goal: Make a short video with someone talking about your goals for this semester. Post this video to your blog with a diary entry.
- Length: At least 3 minutes

Level 3: It's Good For You

Congratulations on getting to Level 3. You are doing great! On the left side of the screen you can see 3 quests. In order for you to get to Level 4, you must complete at

least 2 of the quests. These quests will make you think about other people and challenge you to improve your own life. Good Luck!

Quest 4: Write: A Foreign Friend

- Goal: Have a discussion with a foreign exchange student or any foreign person in Japan and find out about their life.
- Length: At least 150 words
- Instructions: Write about the discussion in your blog. Add photos to make it more interesting. Don't forget to include your weekly diary entry.

Quest 5: Write or Video: Pay it Forward

- 'Pay it forward' is an English expression. It means that you respond to someone's kindness to you by doing nice things for other people.
- Goal: Talk or write about someone who did something nice for you and then talk/write about some nice things you have done/will do for someone else.
- Length:
  * If writing, you should write at least 150 words.
  * If making a video, you should speak for at least 2 minutes.
- Instructions: Post your work to your blog after your weekly diary entry.

Quest 6: Writing: 30-day Challenge

- Goal: Challenge yourself to do something positive for 30 days.
- Length: As long as you like
- Instructions:
  - Watch and study the TED Talk: Try something new for 30 days, in your blog, write down any new or interesting words you hear, and write example sentences related to your life. Then, think of a 30 day challenge you would like to do and write about it in your blog
  - At the end of the 30 days, edit your blog post and write about your experience. Add photos to make your blog more interesting.

Level 4: Looking forward to the future

Welcome to Level 4! These quests are related to your future. On the left side of the screen you can see 3 quests. In order for you to get to Level 5, you must complete at least 2 of the Quests. Good Luck!

Quest 7: Video: Interesting Jobs

- Goal: Make a video with classmates talking about a job you would like to do in the future.
  Length: At least 10 minutes long.
- Instructions: With 1 or 2 other people, make a video interviewing each other and compare jobs. Ask lots of follow up questions just like a conversation. Post this video to your blog after your diary entry.

- Prepare: Before you make this video you should do some research so you can answer the questions

Quest 8: Write: Living Abroad

- Goal: Write about a foreign country you want to live in for 1 year. Add photos/video to make it more interesting. Post this on your blog after your diary entry.
- Length: At least 150 words.

Quest 9: Narrated Presentation: My Major

- Goal: Make a narrated presentation about something related to your major and then post it on your blog after a diary entry.
- Length: At least 5 minutes
- Description: You can choose the topic of your narrated presentation but it should be something related to your major or future work.

Level 5: The World Around You

Level 5 includes video news and TED talks. Both are great tools to study English and learn about the world around you without the need of a teacher. As usual there are 3 quests. In order for you to get to Level 6, you must complete at least 2 of them. Good Luck!

Quest 10: Write: Video News

- Goal: Study and respond to a video news story. Post your response on your blog after your weekly diary entry
- Length: Any length
- Instructions: Watch a video news story on the Reuters website that you are interested in. In your blog, write down any new words you learned. Keep watching the same video until you can completely understand it without looking at the transcript In your blog, answer these 2 questions:
    - What are the main points of the video?
    - What is your opinion about this topic?

Quest 11: Video: Interview

- Goal: Make a video of you interviewing someone in English about a topic you are interested in.
  Length: At least 5 minutes.
- Instructions: Upload the video to your blog after your weekly diary entry.

Quest 12: Speaking & Writing: Ted Talks

- Goal: Watch and study a TED Talks video, then record yourself speaking the same video, using the same words. Length: Any length is fine.

- Instructions: Post this video to your blog after your weekly diary entry. Write any new vocabulary you learned, write about the topic of the talk, and write your opinion about that topic.

Level 6: Learning through new media

You are nearly at the end! Keep going! Let's now look at new media. In Level 6 you can use Vice Media to explore interesting culture, ELLLO to hear world Englishes, and Podcasts to keep your ears and brain busy. These are all great tools to study English by yourself. When you have completed at least 2 of the quests, you can proceed to Level 7. If you are ahead of schedule, you might want to do all 3 quests in this level. Good luck and enjoy.

Quest 13: Video: Documentary Discussion

- Quest: Discuss a VICE documentary in English.
- Length: At least 6 minutes
- Instructions: In a group of 2 or 3 people, choose 1 VICE documentary that you are all interested in. Everyone should watch the documentary. You can watch it together or watch it by yourself. Think of some questions you want to ask your group members about the documentary. In your group, talk about the Vice documentary you all watched. Ask your questions. Record the video. After your weekly diary entry, post the URL of the Vice documentary followed by the discussion video you made. You can write some final comments about the documentary, the discussion you had, or new useful vocabulary you learned in your blog entry too.

Quest 14: Video or Write: ELLLO

- Goal: Watch videos on ELLLO and respond.
- Instructions: You have 2 options for how you can complete this quest. Please choose one of them.
- Option 1: Video: By yourself or in a group of 2-10 people, make your own ELLLO style video. Think of an interesting question or topic to talk about. Any topic is fine. You don't have to write a transcript. It can be as long as you like. 1 person should upload the video to YouTube and tell the other members the URL. Post the video to your blog after your weekly diary entry.
- Option 2: Writing: In your blog, write about 3 videos you watched. Write about the nationalities of the people speaking, their opinions, and your opinion

Quest 15: Listen & Write: Podcasts

- Goal: Find, listen, and review 3 different podcasts that can be used for learning English.
- Instructions: After finding some podcasts you are interesting in, in your blog, write a short review for each podcast. In your review you should write the Podcast name, the Episode name or date, good points, and bad points. Finally, write about which one is the best for you and why.

Level 7: Time to reflect

Congratulations!! You have made it to the final level. You should definitely do Quest 16 and 17. Quest 18 is optional. I hope during these quests you have learnt new things, built stronger relationships with the people around you, and become more confident using English. Your blog is yours... I hope when you look at in the future, it brings back good memories from your time in this class.

Quest 16: Write: Presentation Review

- Goal: Review your presentation
- Length: As long as you like
- Instructions: Ask your teacher for the URL of your presentation. Watch your presentation again. Insert the video of your presentation on to your blog. Reflect on your presentation individually and with your group. Below the video, write the important points which you reflected about.

Quest 17: Write: Reflecting on the semester

- Goal: Write about your semester.
- Length: As long as you like.
- Description: Write about your semester. How was it? You can write about anything you like. You could write about school, work, friendship. Also, answer the questions below:
  1: What was the best thing that happened to you this semester?
  2: Did you achieve your goals? Maybe you talked about them in Quest 3
  3: What do you want to do now that you have finished this English class?

Quest 18: Design your own quest

- Goal: Design your own quest
- Length: As long as you like
- Instructions: Do you have an idea for a good quest? Write about it in your blog. If you have time, you should also do the quest. Anything is fine.

Appendix B: The Language Learning Orientation Scale (LLOS)

Language Learning Orientation Scale

言語習得 (適応) (段階)

The following section contains a number of reasons why one might study a second language. Beside each one of the following statements, write the number from the scale which best indicates the degree to which the stated reason corresponds with one of your reasons for leaning a second language. Remember that there are no right or wrong answers, since people have different opinions.

以下は「なぜ第二言語を学んでいるのか」の幾つかの理由が記載されています。それぞれの文に対し、あなたが第二言語を学んでいる理由の一つに最も当てはまる数字を記入して下さい。みんな違う考え方を持っているので、正解や不正解はありません。

Does not correspond 当てはまらない	Corresponds very little 本当に少しだけ当てはまる	Corresponds a little 少し当てはまる	Corresponds moderately 当てはまる	Corresponds a lot 大変当てはまる	Corresponds almost exactly ほとんど完全に当てはまる	Corresponds exactly 完全あてはまる
1	2	3	4	5	6	7

Why are you studying English? なぜあなたは英語を勉強しているのですか。

Amotivation 動機

1. I cannot come to see why I study English, and frankly, I don't give a damn. なぜ英語を勉強するに至ったのか分からない更に率直に言うと、気にもしていない。
2. Honestly, I don't know; I truly have the impression of wasting my time in studying English. 正直に言うと、分からない。英語を勉強している時間は無駄だという印象を強く持っている。
3. I don't know; I can't come to understand what I am doing studying English. 分からない。なぜ英語を勉強しているのか理解すらできない。

External Regulation 外的（要因）

4. Because I have the impression that it is expected of me. 英語習得は私に期待されている事のような印象を持っているから。
5. In order to get a more prestigious job later on. 後に評価の高い職に就くため。
6. In order to have a better salary later on. 後により良い給与を得るため。

Introjected Regulation（取り入れ　規則）

7. To show myself that I am a good citizen because I can speak English. 英語が話せる事によって自分は良い国民だと示すため。
8. Because I would feel ashamed if I couldn't speak to my friends from English speaking countries in their native tongue. 母国語が英語の友人と話す事が出来ないと恥をかく事になると思うから。
9. Because I would feel guilty if I didn't know English. 英語を知らないと罪悪感を感じる事になると思うから。

Identified Regulation

10. Because I choose to be the kind of person who can speak more than one language. 一つ以上の言語を話せるような人間になりたいと考えたから。
11. Because I think it is good for my personal development. 自己啓発に良いと考えたから。
12. Because I choose to be the kind of person who can speak English. 英語が話せる人間になりたいと考えたから。

Intrinsic Motivation—Accomplishment（本来の動機—達成）

13. For the enjoyment I experience when I grasp a difficult construct in English. 高度な英語の仕組みが把握出来た時の楽しみを経験するため。
14. For the satisfaction I feel when I am in the process of accomplishing difficult exercises in English. 難しい英語の課題を達成する過程での満足感を味わうため。
15. For the pleasure I experience when surpassing myself in my English studies. 英語学習において卓越した自分に対し喜びを経験するため。

Intrinsic Motivation—Knowledge (本来の動機—知識)

16. Because I enjoy the feeling of acquiring knowledge about English speaking communities and their way of life. 英語の世界と彼らの文化の知識を習得する事への感情を楽しんでいるから。
17. For the satisfied feeling I get in finding out new things. 新しい事を発見した時の満足感のため。
18. For the pleasure that I experience in knowing more about English speaking communities and their way of life 英語の世界と彼らの文化をより深く知れた時の喜びのため。

Intrinsic Motivation—Stimulation 本来の動機—刺激

19. For the "high" I feel when hearing English spoken. 英語での会話を聞いた時に感じる高揚感のため。
20. For the "high" feeling that I experience while speaking English. 英語を話している時に経験する高揚感のため。
21. For the pleasure I get from hearing English spoken by native second language speakers. 英語が第一言語の人の英語を理解出来た時の喜びのため。

Appendix C: Leaderboard Questionnaire

Leaderboard Questionnaire リーダーボードについての質問表

Please answer the questions below honestly. I am interested to know how leaderboards affect you. Your data will be kept secure and confidential. How you answer will in no way affect your grade. It should take you about 10 minutes to finish. Thank you for your help! Your time is greatly appreciated. You can use Japanese or English to write your answers. Please use which language you are most comfortable using.
以下の質問に正直に答えて下さい。リーダーボードがどのようにあなたに影響を与えたのかに興味があります。もちろん成績には影響しません。10 分で終わります。ご協力ありがとうございます。※回答データは安全で漏えい又は他の目的に使用することはありません。

You will answer questions about your weekly ranking and your overall ranking. The weekly ranking is only concerned about each week. Not your combined score. For example, during week 3, how did you perform? The overall ranking is your ranking of all the weeks combined.

週刊順位と総合順位について質問に答えてもらいます。週刊順位は各週ごとのみで、すべてのスコアに対してではありません。例）この３週間どのように取り組んだか。総合順位はすべての週の順位を集計したものです。

Weekly Leaderboard 毎週のリーダーボード

1. What was your **weekly** ranking 4 weeks ago? (1-26) あなたの４週間前の順位は？(1−26)
2. What was your **weekly** ranking 3 weeks ago? (1-26) あなたの３週間前の順位は？(1−26)
3. What was your **weekly** ranking 2 weeks ago? (1-26) あなたの２週間前の順位は？(1−26)
4. What was your **weekly** ranking this week? (1-26) あなたの今週の順位は？(1−26)

5. How do you feel when you see your **weekly** leaderboard ranking **this week**? You can choose more than 1 answer if you want to. リーダーボードを見た時どのように感じましたか。(一つ以上選択可)
    a. enjoyment 楽しみ
    b. hope 希望
    c. pride 誇り
    d. determined やる気になった
    e. surprised 驚き

f. anxiety 心配
g. shame 恥かしさ
h. hopelessness 絶望
i. envy 嫉妬
j. Other (please specify) 他?

6. Why did you choose the previous answer? なぜその答えにしましたか？

7. Is the emotion you feel this week, different than the emotion you felt 2, 3, or 4 weeks ago? If yes, please explain why. 今週の感情は 2、 3 、 4 週間前の感情とは違いますか。違う場合は説明して下さい。

Overall Leaderboard 総合のリーダーボード

8. What was your overall ranking 4 weeks ago? (1-26) あなたの 4 週間前の総合順位は？
9. What was your overall ranking 3 weeks ago? (1-26) あなたの 3 週間前の総合順位は？
10. What was your overall ranking 2 weeks ago? (1-26) あなたの 2 週間前の総合順位は？
11. What is your overall ranking this week? (1-26) あなたの今週の総合順位は？

12. How do you feel when you see your **overall** leaderboard ranking **this week**? You can choose more than 1 answer if you want to.リーダーボードを見た時どのように感じましたか。(一つ以上選択可)
    a. enjoyment 楽しみ
    b. hope 希望
    c. pride 誇り
    d. determined やる気になった
    e. surprised 驚き
    f. anxiety 心配
    g. shame 恥かしさ
    h. hopelessness 絶望
    i. envy 嫉妬
    j. Other (please specify) 他?

13. Why did you choose the previous answer? なぜその答えにしましたか？

14. Is the emotion you feel this week, different than the emotion you felt 2, 3, or 4 weeks ago? If yes, please explain why. 今週の感情は 2、 3 、 4 週間前の感情とは違いますか。違う場合は説明して下さい。

15. How hard are you trying to perform well in this class recently? 最近の授業ではどのくらい勉強できているか
    a. 1: I'm not trying at all
    b. 2: I'm trying a little bit
    c. 3: Normal
    d. 4: I'm trying
    e. 5: I'm trying really hard

16. Why did you choose the previous answer? Why are you trying hard or not trying? なぜその答えにしましたか？

Motivation and Effort

How do you feel about the statements below? Please choose the number that corresponds to your feeling 以下の文に対しどう思いますか？（答えに対する適当な数字を選択し記入してください）

1 = strongly disagree (全くそうは思わない), 2 = disagree（そう思わない）, 3 = slightly disagree（どちらかといえばそう思わない）4 = slightly agree（どちらかといえばそう思う）, 5 = agree（そう思う）, 6 = strongly agree（強くそう思う）

17. In relation to leaderboards:
    a. When I see the leaderboard I feel I must do more work to gain rank on the leaderboard.
       リーダーボードを見た時、順位を上げるため更に努力しなければ、と感じる。
    b. When I see the leaderboard I feel I must do work in order to maintain rank on the leaderboard.
       リーダーボードを見た時、順位を維持(いじ)するため努力しなければ、と感じる。
    c. I don’t care where my ranking is on the leaderboard.
       リーダーボードの順位がどこでも気にしない。
    d. When I see the leaderboard I feel I must do work in order to show the teacher that I'm a good student. リーダーボードを見た時、先生に自分は良い生徒だと示すために努力しなければ、と感じる。

e. When I see the leaderboard I feel I must do work in order to show other students that I'm a good student. リーダーボードを見た時、他の生徒に自分は良い生徒だと示すために努力しなければ、と感じる。
f. When I see the leaderboard I feel I must do work in order to prove to myself that I'm a good student. リーダーボードを見た時、自分は良い生徒だと自分自身に証明するために努力しなければ、と感じる。
g. When I see the leaderboard I feel I must do more work to avoid embarrassment.
リーダーボードを見た時、恥をかかないために更に努力をすべきだ、と感じる。

18. In relation to leaderboards:
    a. When I see the leaderboard I feel I must do more work to improve my class grade.
    リーダーボードを見た時、成績を改善するため更に努力しなければ、と感じる。
    b. The leaderboard is fun.
    リーダーボードがあるのは楽しい、と感じる。
    c. The activities (Reading, English Central, Quests) are not enjoyable so I don't care about the leaderboard.
    課題は楽しくなかったので、リーダーボードも気にしない。
    d. When I see the leaderboard I feel, motivated to do more work to improve my English ability.
    リーダーボードを見た時、英語能力を改善しようと動機づけられる。
    e. The leaderboard reminds me that improving my English ability is important.
    リーダーボードボードは、英語能力の改善は重要だという事を思い出させてくれる。

For question 19, 20, 21 and 22. Just choose 1 of them and finish the sentence.
19, 20, 21, 22 の中から一つだけ選び、文を完成させてください。

19. I will work harder because.... 今後は更に努力する、なぜなら～
20. I will continue working at the same rate because ... 今後もこの状態を維持する、なぜなら～
21. I can't work harder because .... 更なる努力はできない、なぜなら～
22. I won't work harder because .... 更なる努力はしない、なぜなら～

23. Generally, what do you think about using Leaderboards in class?
全体的にリーダーボードをクラスで使うことに関してどう思いますか。
24. Do you have any comments about the leaderboards or this class?
リーダーボードに対し何かコメントはありますか？

Thank you for doing this survey.

Leaderboard Questionnaire 3: It is the same as the previous questionnaire until Question 19.

19. This class used some ‘game’ design techniques such as points (30 points for reading, 20 points for English Central, 50 points for each quest), Levels, and Quests. このクラスではポイント制（リーディングに３０ポイント、English Central に２０ポイント、各クエストに５０ポイント）のようなゲーム設計の技法が使われました。

This made the class feel like a game. これが授業をゲームの様にしたのです。

What answer reflects your opinion towards the previous underlined sentence.
上の下線に対するあなたの意見は何ですか。

1: Does not correspond 当てはまらない
2: Corresponds very little 本当に少しだけ当てはまる
3: Corresponds a little 少し当てはまる
4: Corresponds moderately 当てはまる
5: Corresponds a lot 大変当てはまる
6: Corresponds almost exactly ほとんど完全に当てはまる
7: Corresponds exactly 完全あてはまる

* What number (1-7) is your answer? _____

20. Do you have any comments about the ‘game’ aspect of this class? なぜその答えにしましたか？授業のゲームの局面に対して何かコメントはありますか。

21. Generally, what do you think about using Leaderboards in class?
全体的にリーダーボードをクラスで使うことに関してどう思いますか。

22. Do you have any final comments about the leaderboards or this class?
リーダーボードに対し何かコメントはありますか？

Thank you for doing this survey.

Appendix D: General Information Sheet

General Information Sheet for Students English Version

Contact Information
Researcher: Andrew Philpott (andrewphilpott83@gmail.com)
Andrew's Research Supervisor: Jeong-Bae Son (jeong-bae.son@usq.edu.au)

Background

Hello, as you know, my name is Andrew Philpott. I am a lecturer in the Language Center at this university. I am also an external (PhD) doctoral student at the University of Southern Queensland in Australia. I have been in my doctoral course for about 2.5 years. I am researching student motivation related to leaderboards and questing.

Leaderboards are used in English Central and MReader to track and compare your progress with other students. Questing is something new we will do this semester. A quest is an activity you complete using your English skills. The quests for this semester can found here: http://englishquestslevel1.blogspot.jp/
For my doctoral course, I need to collect data about leaderboards and questing. I would greatly appreciate your help for this.

Explanation of class

1. In addition to the normal classroom activities (textbook, TOEIC study, etc.), this semester our Intensive English course will include leaderboards and quests.
2. Each week, your goal is to get as many points (Maximum = 100) as possible by completing three activities for homework: English Central, reading with MReader, and Quests.
3. You can do these activities as much as you like each week (e.g. you can read more than one book); however, your weekly points will be based on the Scoring Table below.

Scoring Table (Weekly Max Points = 100)

Quests (Max 50 points)
- 50 points for successfully completing 1 quest
- 30 points for 50% to 99% quest completion
- 15 points for 1% to 49% quest completion
- 0 points for no quest post on your blog

Reading (Max 30 points)
- 30 points for reading a book and passing quiz
- 15 points for reading a book but failing quiz
- 0 points for no attempt

English Central (Max 20 points) 100% = Watch and Speak 3 videos
- 20 points for 100% completion
- 10 points for 50% - 99% completion
- 5 points for 1%- 49% completion
- 0 points for 0%

4. The weekly deadline for completion is Friday 9:00am.
5. Each week, at the beginning of our computer room class you will randomly be assigned 2 students' quests to assess. The quests will be displayed on their blogs which will also include a personal diary entry.
6. First, you should read the personal diary entry and quest for that week. Then, give their quest a score based on the table above (50, 30, 15, 0 points). Input the scores into the Quest Score Form which is located on our Facebook group.
7. Write a comment in the comments section of their blogs. The comment can be as long as you like. The comment can be about their diary entry or quest.

8. You will now have about 45 minutes of free time to discuss and comment on others students' blogs, work on your next quest, and ask Andrew for help.
9. At the end of this free time you will be shown the class leaderboard which shows your class ranking based on how many points you have accumulated this week.
10. You will be given a couple of minutes to think about and discuss the leaderboard if you like.
11. Every 3 weeks you will be required to do a quick survey. This will be sent to your email. Please open your email and complete the survey. It can be done on your phone or on your computer. It will take you about 5 minutes. Please be honest and thoughtful when answering the questions.

Quest Information

1. Quests are language learning activities.
2. There are 18 quests which you can choose from.
3. The quests are presented in themed levels. 7 levels in total. Level 1 only requires 1 week to be completed; however, the other levels will require 2 weeks to complete.
4. Level 2-7 included 3 quests. When you have completed at least 2 quests from a level, you can progress to the next level. You can do all 3 quests in the level if you like.
5. Different quests will require you to use and develop different language and technical skills.
6. Before you move onto the next level, please do the Quest Diary on the "Change Level" page. Write your opinions about the quests and this class in Japanese or English. Only Andrew will read this.
7. You should do 1 quest a week. However, you can do more if you like.
8. We will use the blogging platform Blogger to display your completed quests. Each week, in 1 blog post, include a personal diary entry before displaying your completed quest.
9. Try to enjoy these quests by working and sharing the experiences with classmates. Be creative and ask for help if you don't understand something.

Benefits to You

1. Improve your English ability and learning skills.
2. Expand your worldview and hopefully increase your motivation to study English.
3. Build stronger relationships with people around you.

Expectations of Participants

1. Complete all the surveys honestly. I want to know your true feelings. Don't be afraid to say negative comments.
2. Treat this as a normal English class. Complete all the work.
3. Ask Andrew if you are not sure about something.

Data collection timetable

For my study to be successful it is important that you complete all the surveys. The table below shows information about the data collection timetable. At the end of the course, some students may be invited to participate in a 30-60 minute focus group discussion.

Survey	Timing	Time required
Motivation survey	week 1 & after week 13	5-10 minutes
Leaderboard survey	Every 3 weeks (4 times total)	About 5 minutes
Quest Diary Survey	At the end of Level 2,3,4,5,6	5-10 minutes
Final Quest Diary Survey	End of Semester	10-20 minutes
Approximate total time required on surveys =		about 1-2 hours

Confidentiality, Privacy, and Consent

1. Any data collected as a part of this project will be stored securely.
2. Personal data will not be shared with anyone not directly involved in this study.
3. Please feel comfortable to share your true opinion when answering the surveys. Your survey answers will not affect your grade in any way.
4. Your participation in this project is entirely voluntary. If you do not wish to take part you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage by contacting a member of the research team.
5. If you do not want to participate in the data collection you will still be expected to do all the same classroom activities.

General Information Sheet for Students Japanese Version

学生用概要シート

連絡先情報
研究者：Andrew Philpott (andrew@kwansei.ac.jp)
Andrew の研究監督: Jeong-Bae Son (jeong-bae.son@usq.edu.au)

経歴

こんにちは、皆様ご存知の通り、私は Andrew Philpott です。この大学の言語教育研究センターで、講師をしています。オーストラリアのサザンクウィーンズランド大学の博士課程学生でもあります。この博士課程には、約２年半在籍しています。私はリーダーボードとクエストに関する学生のモチベーションについて研究しています。リーダーボードはイングリッシュセントラルと M リーダーで、あなたの進歩を追跡し、他の学生達と比較するのに使用されています。クエストは今学期に私達が始める新しいものです。クエストは、英語スキル使って完成させるアクティビティです。今学期のクエストはこのリンクにあります。
http://englishquestslevel1.blogspot.jp/　私の博士課程のために、私はリーダーボードとクエストに関するデータを収集する必要があります。皆様のご協力に、大変感謝します。

クラス説明

１．通常のクラス内アクティビティ（教科書、TOEIC 対策等）に加えて、この学期の集中英語コースには、リーダーボードとクエストが含まれます。
２．各週、あなたの目標は、　課題で３つのアクティビティ：イングリッシュセントラル、M リーダーでの読書、クエスト）を完成し、より多くのポイント（１００点満点）を得ることです。
３．各週、これらのアクティビティを、あなたがしたいだけする事が可能です。（例：1 冊以上の本を読んでも構いません。）しかし、各週のポイントは以下の得点表に基づきます。

得点表

クエスト（５０ポイント）	リーディング（３０ポイント）
・１クエスト無事達成で５０ポイント	・本を読みクイズをパスで３０ポイント
・クエスト５０%～９９%達成で３０ポイント	・本は読んだがクイズは落ちたで１５ポイント
・クエスト１%～４９%達成で１５ポイント	・試みなしは０ポイント
・ブログへのクエスト投稿なしは０ポイント	

イングリッシュセントラル（２０ポイント）１００%＝３つビデオを観て話す

- 100% 達成で 20 ポイント
- 50% - 99% 達成で 10 ポイント
- 1%- 49% 達成で 5 ポイント
- 0%はポイント

４、各週の締め切りは金曜日午前９時です。

５、各週、コンピューター室クラスのはじめに、ランダムで生徒２人分のクエストにアクセスしま す。そのクエストは、個人日記の内容も含んだ彼らのブログに表示されます。

６、まず、その週の個人の日記内容とクエストを読みます、そして、彼らのクエストに上記の表（５０、３０、１５、０ポイント）に基づいたスコアをつけます。そのスコアをフィードバックグループ内のクエストスコアフォームに記入します。

７、彼らのブログのコメント欄に、コメントを記入してください。コメントは好きなだけ書いて構いません。そのコメントは彼らの日記内容かクエストに関するものです。

８、その後、45 分間の自由時間を取ります。この時間は、議論・彼らのブログへのコメント・あなたの次回のクエストへ取り組み・アンドリューへの質問時間です。

９、この自由時間の最後に、今週、何ポイントを獲得したかに基づいた、あなたのクラス内ランキングを示すクラスリーダーボードを公開します。

１０、リーダーボードに関して考え、議論する時間を数分程、取っても構いません。

１１、３週間毎に、短い調査をして頂きます。これはｅメールに送られます。メールを開いて調査を完成させて下さい。携帯やコンピューターで出来ます。5 分程かかります。質問には、本音そしてよく考えてお答え下さい。

クエスト情報

１、 クエストは、語学学習アクティビティです。

２、 １８個のクエストから選択できます。

３、 クエストはテーマレベルごとに提示されています。計７レベルあります。レベル１は完成するのに１週間のみ必要です。しかし他のレベルは完成するのに２週間必要です。

４、 レベル２～７には、各３つのクエストが含まれています。１つのレベルから最低２つのクエストを完成させると、次のレベルに進むことができます。レベル内の全 3 クエストをしても構いません。

５、クエストごとに、異なった言語テクニカルスキルを使用する事になり、それらを改善します。

６、次のレベルに進む前に、“チェンジレベル”ページにあるクエスト日記調査をして下さい。

クエストと、このクラスに関する意見を日本語か英語で書いてください。これはアンドリューのみが読みます。

７、１週間に１クエストはして下さい。しかしもっとしたい場合はして下さい。

８、完成したクエストを表示するのに、ブログプラットフォーム「ブロガー」を使用します。各週、１ブログ投稿、完成したクエストを表示する前に、個人の日記内容を入れて下さい。

９、これらのクエストをクラスメイトと協力し、経験を共有しながら楽しむように心がけてください。創造的になり、何か分からない事があれば助言を求めて下さい。

あなたへの利益

１、あなたの英語能力と学習スキルを改善します。

２、世界観を広げ、上手くいけば英語学習のモチベーションを上げることが出来ます。
３、周囲の人々との強い関係性を築きます。

参加者への期待

１、本音で全調査を完成させて下さい。私は、あなたの本心を知りたいので、ネガティブなコメントを言うことを躊躇わないで下さい。
２、通常の英語クラスとして扱って下さい。すべての作業を成し遂げて下さい。
３、分からない事はアンドリューに聞いて下さい。

データ累積タイムテーブル

私の研究が成功するためには、あなたが全ての調査を終える事が重要です。以下のテーブルはデータ累積タイムテーブルに関する情報を示しています。このコースの最後に、数名の生徒に１５～３０分間のインタビューへの参加をお願いします。

調査	時期	所要時間
モチベーション調査	第 1 週& 13 週以降	5-10 分
リーダーボード調査	毎３週 (計 4 回)	約 5 分
クエスト日記調査	レベル 2,3,4,5,6 の最後	5-10 分
最終クエスト日記調査	学期末	10-20 分
この調査に必要な予測合計時間=		約１～２時間

機密情報、個人情報、同意

１、 このプロジェクトのために収集されたデータは、安全に保存されます。
２、 個人情報は、この研究の関係者以外へは提供されません。
３、 調査に答える際は、安心して本音をシェアして下さい。あなたの調査回答は、成績には一切影響を与えません。
４、 このプロジェクトへの参加は、完全にボランティアです。参加を望まなければしなくて構いません。参加を決めた後に気持ちが変われば、どの段階でも研究チームに連絡後このプロジェクトから辞退する事が出来ます。
５、 このデータ累積に参加したくない場合でも、変わらず同じクラス内アクティビティをして頂きます。

Appendix E: Informed Consent Form

University of Southern Queensland

Consent Form for USQ Research Project Questionnaire

Project Details

Title of Project:	Examining EFL Students' Motivation and Attitudes Toward a Gamified Course Using Leaderboards and Quests at a Japanese University
Human Research Ethics Approval Number:	H16REA038

Research Team Contact Details

Principal Investigator Details
Mr Andrew Philpott
Email: andrew@kwansei.ac.jp

Supervisor Details
Associate Professor Jeong-Bae Son
Email: jeong-bae.son@usq.edu.au
Telephone: +61 7 4631 2235

Statement of Consent 同意書

By signing below, you are indicating that you: 私は署名により、以下の項目に同意致します:

- Have read and understood the information document regarding this project.
  この企画に関する書類を閲覧し、理解しました。
- Have had any questions answered to your satisfaction.
  あなたの期待に応えられる回答を致しました。
- Understand that if you have any additional questions you can contact the research team.
  追加の質問があれば、あなたが調査団体に連絡をとる事を承認致します。
- Understand that you are free to withdraw at any time, without comment or penalty.
  許可や罰則なしに、いつでも情報を使用する事を承認します。
- Understand that unidentifiable data collected in this study could be used in the future by researchers interested in any issues related to this study. 発信源不明の情報はこの研究に集められ、将来、この研究に関係した発行物に興味をもった調査者に使われる可能性があります。
- Understand that you can contact the University of Southern Queensland Ethics Coordinator on (07) 4631 2690 or email ethics@usq.edu.au if you do have any concern or complaint about the ethical conduct of this project. 気がかりになる事や苦情があれば、USQ EC(電話(07)4631 2690 メール ethics@usq.edu.au)に連絡を取ることを承認致します。
- Are over 18 years of age. 18 歳以上です。
- Agree to participate in the project. この企画に参加する事に同意いたします。

Participant Name 参加者 氏名	
Participant Signature 参加者 署名	
Date 日付	

What do you want your nickname to be on the leaderboard? _____

When completed by the researcher, would you like a summary of the results sent to you? yes / no

Please return this sheet to a Research Team member prior to undertaking the questionnaire.

Appendix F: Semi-Structured Interview Summaries

Semi-structured Interview Summary 1

Participant # 1	Gender female	Final LB rank = 1 category = high	LLOS FL Motivation = extrinsic* ext pre = 2.55 int pre = 1.77 ext post = 1.77 int post = 2.11

Quest summary

1. At the beginning did writing quests because they could be completed quickly.
2. Was initially reluctant to do quests that involved making a video, but after recording a video, realised it was fun and was able to enjoy studying English.
3. Enjoyed the social aspect of questing: collaborating with classmates to make a video, interview, and discussion style quests; seeing how classmates completed quests.
4. Enjoyed meaningful quests such as interviewing a friend about her experience living abroad.
5. Did not enjoy the long writing quests as much as the other quests.

Leaderboard summary

1. Motivated to keep leaderboard ranking.
2. Thinks leaderboards are good because if people see that others are working hard then they will think they should work hard too.

Semi-structured Interview Summary 2

Participant # 2	Gender female	Final LB rank = 16 category = middle	LLOS FL Motivation = intrinsic* ext pre = 1.00 int pre = 3.22 ext post = 1.88 int post = 3.66

Quest summary

1. Quests were fun, especially recording videos with friends or family.
2. Tried hard to make funny videos that classmates would enjoy.

Leaderboard summary

1. leaderboards were initially fun but then became less fun and became more pressure.
2. Influenced everyone to do homework because classmates can see everyone's ranking.
3. Felt like she should have worked harder when she did not get the weekly full score.

Semi-structured Interview Summary 3

Participant 6	Gender female	Final LB rank = 1 category = high	LLOS FL Motivation = mix* ext pre = 3.66 int pre = 3.55 ext post = 3.22 int post = 4.00

Quest summary

1. Quests were hard when she did not have much time, but she appreciated being able to choose quests depending on how busy her weekly schedule was.
2. Completing one quest a week was a good frequency.
3. Good opportunity to practice speaking English and learn news way to study English.
4. Did not do quests that required recording a video because she was shy, but she did enjoy collaborating with classmates.

Leaderboard summary

1. When she looked at the leaderboard she thought that she should work hard to complete all her homework.
2. Thought the leaderboard was good to keep everyone motivated. If there was no leaderboard, maybe she would have sometimes not done some weekly homework.
3. Wanted to maintain her 1st place ranking.

Semi-structured Interview Summary 4

Participant 12	Gender female	Final LB rank = 1 category = high	LLOS FL Motivation = mix* ext pre = 2.88 int pre = 3.55 ext post = 4.00 int post = 6.44

Quest summary

1. Initially, quests were technically difficult. However, she learned new PC skills and could complete quests faster as the semester progressed.
2. Did not want to record videos of herself talking because she felt ashamed.
3. Preferred quests that she could completed by herself such as writing quests. She said she was busy and did not have time to work with classmates.
4. Appreciated the quests that made her reflect on her life because she does not normally think about things like ‘why do you study English?’

Leaderboard summary

1. Cared more about getting maximum points than ranking.
2. She could confirm that she did the homework correctly.
3. The weekly leaderboard deadline made her feel like she should do the homework.

Semi-structured Interview Summary 5

Participant 13	Gender female	Final LB rank = 1 category = high	LLOS FL Motivation = mix* ext pre = 4.66 int pre = 5.11 ext post = 4.88 int post = 5.55

Quest summary

1. Quests were fun because she could be creative. She enjoyed making videos and narrated PowerPoint presentations.
2. Initially quests were technically difficult, but she got used to them and learned various things.

Leaderboard summary

1. Thought leaderboards were good because they motivated her to complete all aspects of her homework.
2. The weekly schedule of homework was sometimes hard. Some weeks she was busy, other weeks she was not busy.

Semi-structured Interview Summary 6

Participant 17	Gender female	Final LB rank = 19 category = low	LLOS FL Motivation = extrinsic* ext pre = 5.11 int pre = 4.11 ext post = 4.11 int post = 2.66

Quest summary

1. Initially quests were difficult but after a while became used to them.
2. Quests were technically difficult, but she was happy when people around her helped her.
3. Enjoyed easy and fun quests such as watching videos, making videos with classmates, and writing about what country she wanted to visit.
4. Quests were a heavy workload around the end of semester due to many assignments from other classes.

Leaderboard summary

1. She said that leaderboard rankings showed basically that you did everything or nothing. Because even if you did not do one small part, your ranking would drop low because most people did everything and got 1st ranking.
2. Couldn't improve overall ranking to become 1st, could only get first in the weekly leaderboard. Needs to be a way to catch up.
3. Didn't think the point system was meaningful enough. Thought the point system needs to more accurately represent the amount and quality of work, rather than just doing the work.

Semi-structured Interview Summary 7

Participant 19	Gender male	Final LB rank = 25 category = low	LLOS FL Motivation = mix* ext pre = 3.77 int pre = 4.11 ext post = 3.55 int post = 3.22

Quest summary

1. There was a big discrepancy in the time required to complete different quests.
2. Didn't like quests that took a long time to make such as the narrated PowerPoint presentation quests.
3. Liked listening and watching videos then stating opinion.
4. He chose quests he could do by himself because he was busy.

Leaderboard summary

1. When he saw his leaderboard ranking, he thought he should do more work, but he didn't.
2. Was motivated to complete the quests, but because he could not catch up to everyone on the leaderboard, he wasn't motivated by it.
3. Enjoyed the leaderboard in the beginning but as his ranking dropped, the more he didn't want to see it, the more he lost interest in the leaderboard.

Semi-structured Interview Summary 8

Participant 22	Gender male	Final LB rank = 1 category = high	LLOS FL Motivation = extrinsic* ext pre = 4.44 int pre = 1.77 ext post = 3.11 int post = 1.77

Quest summary

1. Initially difficult but became easier to understand the quest system as the semester progressed. It became work to him after a while.
2. Appreciated learning new ways to study English.
3. Good opportunity to practice writing and speaking in English. He liked quests where he could have critical conversations with groups of people.
4. Liked easy quests such as writing quests because he could do them on the train and then post his work on his blog easily.
5. He felt that quests that took a long time, such as making videos, were troublesome.

Leaderboard summary

1. He said he's a serious person, he pressure himself, felt like it was his duty to be working hard, making sure he completed everything.
2. There were many students who had the number 1 ranking. He wanted to be the only person to be number 1. He wished that he could have done more work to become the only number 1 ranked participant. Because he couldn't do any extra work, all he could do was wait for someone to make a mistake.
3. Sometimes him and his classmates made fun of students who were low on leaderboard. He said, as the leaderboard ranking did not reflect English skill, just if you did the work on not, some students took it light-heartedly.

Semi-structured Interview Summary 9

Participant 26	Gender male	Final LB rank = 17 category = middle	LLOS FL Motivation = intrinsic* ext pre = 3.33 int pre = 4.44 ext post = 2.77 int post = 4.22

Quest summary

1. Quests seemed interesting, but he was too busy. Because he was busy, they seemed troublesome in the end. He wanted to try recording a video but didn't have time. He just chose quests he could complete quickly.
2. He wished there was time in class to do quests.
3. He enjoyed doing the quest in which he could speak with his foreign friend and then write about it on his blog.

Leaderboard summary

1. He was apathetic towards the leaderboards, didn't care about his middle ranking, because he said he was too busy and did not care about his grade.
2. He thinks the leaderboard is a good reference to see how much work he is doing compared to other participants. It shows who is working hard and rewards those who are high on the leaderboard.
3. Some students talked about the students with low ranking.