Single-player RPGs as a medium of instruction in formal foreign language education

正規外国語教育における教育媒体としてのシングルプレイヤーRPG

Kevin Reay Wrobetz

Abstract

Since the integration of computers into formal academic environments, digital games have been hypothesized to have the pedagogical potential to facilitate the learning process. Propelled by the commercial expansion of the digital game market, the range of pedagogical applications in formal academic environments for digital games is vast and encompasses many subjects from mathematics to second language acquisition (SLA). Research in the field of computer-assisted language learning (CALL) regarding the efficacy of game-based language learning has proposed that massively multiplayer online role-playing games (MMORPGs) may facilitate SLA through a sociocultural perspective of language learners interacting with linguistically more capable peers within the context of the game. Research has yet to identify whether the benefits to SLA observed in MMORPGs is solely due to the interaction between players or perhaps due in part to the gameplay mechanics of role-playing games (RPGs). In order to address the question of whether the gameplay mechanics of digital RPGs are effective at facilitating SLA without the inter-player cooperation seen with MMORPGs, the smartphone edition of the single-player RPG *Life Is Strange* (Dontnod Entertainment, Decknine, Square Enix, 2015; https://square-enix-games.com/en_US/games/life-is-strange) was implemented into a mixed-methods study. The research was conducted from April 2019 to January 2022 at a Japanese university and involved 172 students of mixed English proficiency ranging from novice-low to advanced-high using the American Council on the Teaching of Foreign Languages (ACTFL) proficiency guidelines (Breiner-Sanders, Lowe, Miles, & Swender, 2000). The study's primary aim was to measure the efficacy of target language (TL) vocabulary acquisition and retention in an intervention group (i.e., students who played the game) relative to a control group (i.e., students who did not play the game) with a five-week testing protocol which included vocabulary pre-tests, vocabulary tests issued immediately after the intervention method (i.e., treatment or control) but prior to corrective feedback, delayed vocabulary tests after corrective feedback, and one-semester delay vocabulary tests. Furthermore, the study aimed to identify what factors affected vocabulary acquisition during gameplay, the students' perceptions of using RPGs as a language-learning resource, and how the game content influenced cultural awareness of the TL with comprehension tests based on in-game content, pre-/post-game surveys, and in-class student projects based on the efficacy of digital game-based learning in the classroom. The results suggest that single-player RPGs not only effectively elicit long-term vocabulary acquisition but may also increase students' perceptions of smartphone gaming as an effective language learning tool and influence the perception of cultural aspects of the TL.

概要

コンピュータが正規の教育環境に導入されて以来、デジタルゲームは学習プロセスを促進する教育的可能性を持っているという仮説が立てられてきました。デジタルゲーム市場の商業的拡大に伴い、デジタルゲームの教育的応用範囲は膨大で、数学から第二言語習得（SLA）まで多くの教科に及んでいます。コンピュータ支援型言語学習（CALL）の分野では、ゲームベースの言語学習の効果に関する研究により、多人数参加型オンライン・ロールプレイングゲーム（MMORPG）は、言語学習者がゲーム内で言語能力の高い仲間と交流するという社会文化的視点を通じて、SLA を促進する可能性があることが提唱されています。MMORPG で観察される SLA へのメリットが、プレイヤー間の相互作用のみによるものか、あるいはロールプレイングゲーム（RPG）のゲームプレイメカニクスに一部起因するものなのかは、まだ研究によって明らかにされていません。デジタル RPG のゲームプレイメカニクスは、MMORPG で見られるようなプレイヤー間の協力関係なしに SLA を促進する効果があるのか、という疑問を解決するために、シングルプレイヤーRPG『Life Is Strange』（Dontnod Entertainment, Decknine, Square Enix, 2015; https://square-enix-games.com/en_US/games/life-is-strange）のスマホ版を混合法研究に導入しました。研究は、2019 年 4 月から 2022 年 1 月まで日本の大学で行われ、アメリカ外国語教育協議会（ACTFL）の習熟度ガイドライン（Breiner-Sanders, Lowe, Miles, & Swender, 2000）を用いて初級-下から上級-上までの英語力が混合した学生 172 人が参加しました。本研究の主要目的は、介入群（ゲームを行った学生）が対照群（ゲームを行わなかった学生）と比較して、目標言語（TL）語彙の習得と定着の効果を測定することであり、語彙の事前テスト、介入方法（すなわち、介入または対照）直後の修正フィードバック前の語彙テスト、修正フィードバック後の遅延語彙テスト、1 学期遅延語彙テストなどの 5 週間のテストプロトコルを実施しました。さらに、ゲームプレイ中の語彙習得に影響を与える要因は何か、RPG を言語学習教材として利用することに対する学生の認識、ゲーム内容が TL の文化認識にどのような影響を与えるかを、ゲーム内コンテンツに基づく理解テスト、ゲーム前後のアンケート、授業におけるデジタルゲーム学習の有効性に基づく学生プロジェクトで明らかにすることを目的としました。その結果、シングルプレイヤーRPG は長期的な語彙習得を効果的に引き出すだけではなく、スマホゲームを効果的な言語学習ツールとして学生の認識を高め、TL の文化的側面に対する認識に影響を与える可能性が示唆されました。

Acknowledgements

Professionally, I would like to thank to my PhD advisor, Dr. Mark Peterson, for providing guidance and feedback throughout this project. I would also like to thank all my colleagues in Mark Peterson's research lab whose research projects have helped refine my own research methods. Finally, I would like to thank all my past teachers and mentors whose encouragement has given me the motivation to pursue an academic career.

Personally, I would like to thank all my family and friends who have encouraged me to pursue my interests. Without their support, I would not have been able to transform my hobbies into my job. Above all, I would like to express my deepest gratitude to my wife, Orie Wrobetz, and my sons, June and Reay, whose unconditional love and support drive me to keep pushing forward.

Table of contents

List of figures

List of tables

List of appendices

Glossary of terms

ANOVA: analysis of variance

CALL: computer-assisted language learning

CF: corrective feedback

DGBL: digital game-based learning

EFL: English as a foreign language

GS: gaming session

L1: first language

L2: second language

LC group: LIS control group (i.e., just the vocabulary tests)

LG group: LIS game group (i.e., game and vocabulary/comprehension tests)

LIS: *Life Is Strange* (Dontnod Entertainment, Decknine, Square Enix 2015)

MMORPGs: massively multiplayer online role-playing games

NPC: non-playable character

OSD: one-semester delay

PC: personal computer

PS group: pilot study group (i.e., just the game)

RPGs: role-playing games

RQ: research question

SLA: second language acquisition

TBLT: task-based language teaching

TL: target language

Chapter 1 Introduction

Digital games have increasingly garnered attention within the field of computer-assisted language learning (CALL) due in part to intrinsic design qualities which enable them to motivate players to keep playing (Lohse, Shirzad, Verster, Hodges, & Van der Loos, 2013; Takatalo, Häkkinen, Kaistinen, & Nyman, 2010; Tychsen, Hitchens, & Brolund, 2008), educate on how to play while playing (Gee, 2003a; Turkay, Hoffman, Kinzer, Chantes, & Vicari, 2014), and utilize game-specific knowledge to progress through incremental increases in difficulty (Aponte, Levieux, & Natkin, 2009; Gee, 2003a). Although these design features common to many commercially successful digital games may not have a direct relationship with the pedagogical processes of second language acquisition (SLA), when digital games are played in the target language (TL), these intrinsic qualities of motivation, education, and application of knowledge have resulted in TL gains (Piirainen–Marsh & Tainio, 2009; Reinders & Wattana, 2012; Sylvén & Sundqvist, 2012). Considering the breadth of commercial digital games available which may be used for SLA purposes, many aspects of how digital games elicit SLA remain unanswered in relevant literature. Specifically, issues regarding the most efficacious digital game genres, platforms, and pedagogical implementations still require elaboration to fully realize the true pedagogical potential of digital games in SLA contexts. The focus of the present thesis is to explore the potential pedagogical impact of smartphone single-player role-playing games (RPGs) on specific SLA processes in formal educational environments.

The digital game that was chosen as the primary subject of analysis is the single-player, choose-your-own-adventure role-playing game (RPG) *Life Is Strange* (henceforth LIS) (Dontnod Entertainment, Decknine, Square Enix 2015), a genre that is both underrepresented in digital game-based learning (DGBL) and CALL literature and potentially as pedagogically effective as other digital game genres theoretically conducive to SLA, namely massively multiplayer online role-playing games (MMORPGs), virtual worlds, and simulation games (Cornillie, Thorne, & Desmet, 2012; Peterson, 2013; Ranalli, 2008; Sylvén & Sundqvist, 2012; Zhao & Lai, 2009). The smartphone application edition of LIS was chosen as the gaming platform due to its potential to lead to more successful pedagogical outcomes in formal educational environments by way of its scalability, low cost, universal penetration of relevant hardware, and mobility allowing for student agency over the learning process (Wrobetz, 2021). This thesis specifically investigates through mixed-method analysis how the integration of the smartphone application of LIS into English as a foreign language (EFL) courses at a Japanese university affected second language vocabulary acquisition and retention, the perception of digital games to function as a medium of EFL instruction, and the perception of the TL culture. Furthermore, this thesis identifies predictive factors through correlational analysis that may affect pedagogical outcomes of curricula that adopt DGBL as a primary method of instruction in the foreign language classroom.

1.1 Research context and goals

The primary goal of this study was to investigate the efficacy of using single-player RPGs on smartphone gaming platforms as a medium of instruction within the formal academic setting of university-level EFL classrooms. This study specifically analyzed the efficacy of the LIS gaming intervention by measuring the acquisition and long-term retention of in-game TL vocabulary relative to a control group with no exposure to the LIS gaming intervention. Other goals of the present study were to identify correlational factors between vocabulary acquisition and retention outcomes from in-class gaming sessions, measure shifts in attitudes towards English smartphone gaming, and gauge how much the students' perceptions of US culture were influenced by the narrative content of LIS. The course of the present study was structured with the above research goals in mind.

The following chapter will outline the specific relevance of this study through a review of concepts discussed in relevant literature regarding the efficacy of digital games in the field of CALL and detail the relationship between digital games and pedagogical practices in formal educational environments, specifically within foreign language classrooms. Furthermore, a historical account of the evolution of digital games into commercially viable products will be provided to draw attention to how the selection of digital game titles and platforms may influence pedagogical outcomes. These considerations for designing DGBL curricula are then discussed within the context of foreign language education and frame the research goals and questions of the present study.

Chapter 2 Review of the literature

2.1 Introduction

The following chapter will highlight some of the issues with designing and developing DGBL curricula in formal educational contexts as well as discuss how research in the field of CALL has focused on the pedagogical potential of commercially developed digital games. Furthermore, the chapter will discuss how the gameplay mechanics of commercial RPGs on smartphone gaming platforms may be particularly well suited for TL vocabulary acquisition and retention within DGBL curricula. In so doing, this chapter will shed light on the relevance of and clearly state the rationale for using the smartphone edition of LIS for EFL vocabulary acquisition and retention within formal foreign language educational contexts.

2.1.1 Issues with integrating DGBL curricula in the classroom

The potential of DGBL to lead to successful pedagogical outcomes notwithstanding, it may be argued that the success of a DGBL curriculum is largely dependent on the efforts of the instructor to utilize the content of the game as the basis of instruction; put another way, successful pedagogical outcomes of DGBL curricula are not an automatic product of simply sitting students in front of a computer and walking away (All, Castellar, & Van Looy, 2016; Cornillie, Clarebout, & Desmet 2012; Miller & Hegelheimer, 2006; Ranalli, 2008). Just as with any medium of instruction, traditional or otherwise, much of the literature on pedagogical theory points to the necessity for instructors to incorporate instructional media into an active dialogue between the students and the subject of study (Anderson, 2003; Doolittle, 1997; Krashen, 1985; Long, 1980), so too must digital games become an integral part of an ongoing in-class dialogue for the salient pedagogical functionality of the game to come to the forefront. This need for continued dialogue with the digital game may therefore be incongruent with standardized curricula which must also devote in-class time to areas of instruction unrelated to the content of a DGBL curriculum.

2.1.2 Commercial influence on the design of digital games

As digital games became commercially viable products through the continued market penetration of relevant hardware such as personal computers and console gaming systems such as the Nintendo Entertainment System (Marchand & Hennig-Thurau, 2013), the development process became more focused on ludic gameplay mechanics. In principle, the success of commercial digital games hinges entirely on how entertaining the game is to play. The commercial viability principle which drove the evolution of most digital games has been quite successful at creating a wide array of captivating digital game titles as well as a massively affluent subsection of the entertainment industry, the total value of which is estimated at $159.3 billion in 2020 (WePC, 2021), larger than both the movie industry and the music industry combined (Saltzman, 2021). The point must be raised, however, that although the ludic elements of digital games influence the pedagogical potential of DGBL curricula, this does not entail that all ludic gameplay mechanics are particularly well suited for formal educational environments.

The most logical approach to successfully incorporate a commercial digital game into a DGBL curriculum is through a careful selection process. Fortunately for educators working in the field of DGBL, the rapid expansion of the commercial market for digital games has created a large selection of pedagogically viable instructional media. An ideal candidate to serve as a medium of instruction in a formal educational environment would be a digital game in which the ludic gameplay mechanics and narrative content are also able to impart some fundamental principle of the underlying curriculum (Gee, 2005; Habgood & Ainsworth, 2011; Li, 2019). Furthermore, digital games which are accessible outside of the classroom environment, whereby students can exercise their own agency over the learning process, should be prioritized (Shaffer & Gee, 2006). A selection process as outlined above would therefore require careful selection of the narrative content, gameplay mechanics, and gaming platform for educators to maximize the pedagogical potential of incorporating commercial digital games into formal educational environments.

2.1.3 Game selection for the foreign language classroom

Not all commercial digital games are created equal for the purpose of serving as the primary medium of instruction in formal educational environments, especially considering that some research has identified that the inherent interactivity of games may interfere with the language learning process (DeHaan, Reed, & Kuwada, 2010; Klimova & Kacet, 2017). Researchers in the field of DGBL often make the distinction between pedagogically viable and nonviable games through the title of "serious games," or in other words, the difference between digital games with pedagogical potential and those with little to no application in educational contexts (Breuer & Bente, 2010; Connolly et al., 2012; Susi, Johannesson, & Backlund, 2007). Although there is no well-defined rule as to what the gameplay mechanics of serious games must engender in order to qualify for the title, establishing an index for the learning potential of gameplay mechanics and their applicability to out-of-game contexts may serve as a general selection principle. Ultimately, the "seriousness" of any digital game title for educational instruction depends on the underlying goals of the curriculum, the ability of the gameplay mechanics to elicit this process, and the applicability of what is learned to out-of-game contexts. With these selection principles in mind, the question of what a serious game entails for the purpose of SLA is summarized in Table 2.1.

Table 2.1 Selection principles for serious digital games in SLA

Active linguistic skills addressed through gameplay			
Speaking	Genre examples	Writing	Genre examples
Promotes in-game inter-player communication in the TL (with microphone)	MMORPGs, strategy games, first-person shooters	Promotes in-game inter-player communication in the TL (with chat box)	MMORPGs, strategy games
Enables multiplayer functionality to elicit TL communication	Fighting games, action games, first-person shooters, sports games	Enables optional TL script choices in response to in-game dialogue/scripts	MMORPGs, single-player RPGs, puzzle games

Passive linguistic skills addressed through gameplay			
Reading	Genre examples	Listening	Genre examples
Encourages engagement with in-game descriptions or game rules in the TL	MMORPGs, single-player RPGs, strategy games, puzzle games, simulation games	Necessitates the comprehension of inter-player conversations in the TL	MMORPGs, first-person shooters
Necessitates TL content comprehension of in-game narrative content to progress through the game	MMORPGs, single-player RPGs, action games, simulation games, strategy games	Necessitates the comprehension of voice-acted narrative content/dialogue in the TL	MMORPGs, single-player RPGs, action games

Note. The genre examples provided above do not constitute a comprehensive list

With the selection principles outlined above in mind, the question of what pedagogically viable gameplay content/mechanics entail for the purpose of SLA must now be addressed. Simulation and role-playing games have been shown to expose students to authentic use of the TL (Hitosugi, Schmidt, & Hayashi, 2014; Hubbard, 1991; Peterson, 2010a). Role-playing games, virtual environments, simulation games, computer games, and educational games have been shown to benefit TL vocabulary acquisition and retention of in-game vocabulary through repeated exposure (Bytheway, 2014; Lai, Ni, & Zhao, 2012; Li, 2019; Miller & Hegelheimer, 2006; Ranalli, 2008), spaced presentation of TL vocabulary (Clark, Tanner-Smith, & Killingsworth, 2016; Nietfeld, 2017), and multi-modal presentation of vocabulary (e.g., simultaneous representation of TL vocabulary aurally, visually, and textually) (Aghlara & Tamjid, 2011; Zou, Huang, & Xie, 2021). Simulation games, narrative-based games, and role-playing games have been shown to provide corrective feedback or otherwise require comprehension of linguistic input to progress through the game (Chen, Chen, & Dai, 2018; García-Carbonell, Rising, Montero, & Watts, 2001; Reinhardt & Thorne, 2016). Action, role-playing, simulation, strategy, and puzzle games have been shown to create conducive learning environments from a sociocultural perspective through incremental increases in difficulty and interaction with more-capable peers (Gee, 2003b; Hung, Kuo, Sun, & Yu, 2013). Role-playing, action, simulation, fighting, sports, and strategy games have been shown to increase student agency over the learning process through extramural engagement (Kirriemuir & McFarlane, 2004; Sylvén & Sundqvist, 2012). Role-playing games, strategy games, action games, and virtual environments have been shown to encourage engagement with game-based discourse communities (Squire, 2006; Sykes & Dubreil, 2019). Role-playing, action, and quiz games have been shown to reduce anxiety associated with linguistic production in the TL through virtual avatars (Reinders & Wattana, 2014) and increase levels of motivation to engage with the TL (Birk et al., 2016; Wichadee & Pattanapichet, 2018).

The observed benefits that DGBL can have on SLA processes as outlined above are not universal and the degree to which these benefits may have an effect depend on both the content of the game as well as on the gameplay mechanics. A puzzle game that does not make use of voice-acted content such as *Tetris* (Pajitnov & Pokhilko, 1984), for example, would certainly not be as effective for TL vocabulary acquisition as a puzzle game such as *Tangle Tower* (SFB Games, 2019) that does. Similarly, an action game that does not require comprehension of the linguistic input in order to progress through the game such as *Tekken* (Bandai Namco Studios, 1994) may not be as effective for SLA purposes as an action game such as *Grand Theft Auto: San Andreas* (Rockstar North, 2004) that does. Apart from the basic genre of a digital game, what the game is about and how the player progresses through the game can also impact how suited any given digital game is to a DGBL curriculum.

These distinctions, as outlined above, of how gameplay content/mechanics can influence their efficacy as instructional media in SLA contexts must be given particular scrutiny in order to select the most appropriate (e.g., the most “serious”) digital game titles for foreign language classrooms. Social digital games such as MMORPGs have been routinely identified in DGBL literature as being particularly beneficial in SLA contexts due to their innate ability to address both active and passive linguistic skills (Autio & Takamaa, 2020). However, when considering how gameplay content/mechanics may also impact the efficacy of the DGBL curriculum, then other genres of digital games, even those that may not train active linguistic skills, must also be considered based upon how conducive their individual gameplay content/mechanics are at eliciting long-term, meaningful gains in the TL.

2.1.4 Gaming platform selection for DGBL curricula

The final crucial factor that must be considered when implementing a DGBL curriculum is how the gaming platform may affect the pedagogical outcome of the course. The primary gaming platforms for digital games are console systems (e.g., Nintendo Switch, Sony PlayStation), computers (e.g., PCs, laptops), and smartphones. The choice between these platforms affects a number of factors relating to accessibility, cost, scalability, gaming experience, and potential for engagement (Wrobetz, 2021). Table 2.2 outlines some of the advantages and disadvantages for each gaming platform. The following section will briefly address how the advantages and disadvantages of each gaming platform may affect pedagogical outcomes regarding formal implementation in educational contexts.

Table 2.2 Gaming platform advantages and disadvantages

Gaming platform	Pedagogical advantages	Pedagogical disadvantages
Console systems	Widespread market penetration with 80% of households owning a console as of 2014 (North, 2015) and a 25% share of the total gaming market (WePC, 2021) may encourage extramural engagement with relevant hardware/software used in a DGBL curriculum.	High levels of market diversity and competition between console gaming system manufacturers (i.e., Nintendo, Sony, and Microsoft) create an incentive for in-house development of software and copyright ownership (Robert, 2012; Thomes, 2015). The market diversity may increase the costs associated with procuring the necessary hardware to run specific software in DGBL curricula.

PCs	Widespread market penetration with 80% of households owning a PC (Alsop, 2021) and a 24% share of the total gaming market may encourage extramural engagement with relevant hardware/software used in a DGBL curriculum. Widespread access to PCs at academic institutions in industrialized nations may reduce the costs associated with DGBL curricula.	High levels of market diversity in the gaming industry and competition between operating systems (e.g., Microsoft and Macintosh) incentivize exclusive access to digital game titles on one operating system. Access to PCs is not guaranteed at all academic institutions. Further, limited supplies of costly PC hardware may limit the amount of time that one class can spend in computer labs as multiple classes may require access to the PCs at the same time.
Smartphones	Widespread market penetration with 85% of the US population owning a smartphone (Pew Research Center, 2021) and 50% share of the total gaming market (Clement, 2021) may encourage extramural engagement with relevant hardware/software used in a DGBL curriculum. As of 2016, smartphones were the most widely used gaming platform at 71%, PCs at 64%, tablets at 34%, and gaming consoles at 26% (Hutchinson, 2016) which may further increase extramural engagement with a DGBL curriculum. The mobility of smartphones may increase extramural engagement and ensure that all students have in-class access to relevant hardware/software. Widespread market penetration shifts the costs onto students, which significantly reduces DGBL curricular costs and enables DGBL curricula to scale up to larger class sizes.	High levels of market diversity in the gaming industry and competition between smartphone operating systems with Google's Android at 72.93% and Apple's iOS at 26.53% of the global market share (StatCounter, 2021) incentivize exclusive access to digital game titles on one operating system. Smaller screen sizes and limited controller interfaces may increase the difficulty to interact with more complex digital games, which may in turn reduce the total number of pedagogically efficacious games to use in a DGBL curriculum.

Console gaming systems, PCs, and smartphones are all viable gaming platforms for a DGBL curriculum. However, smartphones present unique characteristics which may allow for higher levels of engagement both in and outside the classroom (Moghaddas & Bashirnezhad, 2016; Sharples, Taylor, & Vavoula, 2005; Stockwell, 2010; Stockwell & Hubbard, 2013). Console gaming systems, while potentially available outside the classroom, are not typically available within academic institutions making them a less than ideal option for a DGBL curriculum in a formal educational context (Wrobetz, 2021). In contrast, PCs are usually available in limited capacities within academic institutions and, owing to high levels of market penetration, may also allow access to the digital game being used as the primary medium of instruction outside the classroom. However, DGBL curricula that make use of

PCs as their primary gaming platform are limited by the availability of the PCs at academic institutions as well as by financial constraints, which may create barriers to implementing such curricula (Tüzün, 2007; Wrobetz, 2021). Smartphones are perhaps the most ideal candidate to serve as the primary gaming platform for a DGBL curriculum in a formal educational context due to their high level of market penetration, ability to provide access to the digital game both in and outside the class, and overall scalability (Wrobetz, 2021). With the selection criteria of both game and platform as described above in mind, smartphone-based RPGs may benefit SLA in formal educational contexts. Given the variety of digital titles to choose from within the wider RPG genre, the following section will introduce the key characteristics of MMORPGs and single-player RPGs as they relate to SLA contexts.

2.2 MMORPGs vs. single-player RPGs in SLA contexts

This thesis adopts an interactionist stance of SLA, whereby TL interaction with digital games is hypothesized to elicit linguistic gains through several distinct mechanisms. The types of TL interaction potentially beneficial to SLA in/with digital games can range from interactions between the language learner and the gaming platform/mechanics as well as inter-player social interactions elicited by the gaming platform/mechanics. Although highly dependent on the platform and design of the gameplay mechanics, several psycholinguistic benefits to SLA such as in-game scaffolding (i.e., gradual increases in difficulty), in-game corrective feedback/negotiation of meaning of in-game objectives, contextualization of the TL used in in-game objectives, and long-term retention of TL vocabulary through such mechanisms as spaced/multimodal presentation may be at play when language learners interact with digital games. From a sociocultural perspective, digital games may elicit SLA by embedding language learners in a network of linguistically more capable peers, whereby gains in the TL may be mediated by zones of proximal development, creating TL dialogue mediated by shared knowledge of the gameplay content/mechanics, and establishing intersubjectivity created by the shared gaming experience. The following sections will use relevant CALL literature to outline how both MMORPGs and single-player RPGs may benefit SLA in formal educational contexts.

2.2.1 MMORPGs and social gaming

Returning to the topic of which types of digital games and which specific gameplay mechanics constitute the most beneficial for SLA (see Table 2.1), the merits of MMORPGs have been continually lauded in relevant CALL and DGBL research (Li, 2019; Peterson, 2010a; Peterson, 2012a; Peterson, 2016; Rankin, Morrison, McNeal, Gooch, & Shute, 2009). As to the specific mechanisms at work that aid in SLA processes through engagement with MMORPGs, the social elements of gameplay are well documented in the literature and functionally allow language learners playing MMORPGs in the TL to receive authentic TL input from linguistically more capable peers (Peterson, 2012b; Sylvén & Sundqvist, 2012). MMORPGs also create environments conducive to corrective feedback through both the gameplay mechanics and social network, (Chen, 2015; Goh, 2016; Li, 2019; Peterson, 2010a; Peterson, 2016). Furthermore, the gameplay mechanics and social network provide language learners the opportunity to negotiate the meaning of unknown vocabulary/grammatical structures (Dixon, 2104; Peterson, 2012a, Peterson, 2016). Finally, the virtual environment of pseudo-anonymity may help to alleviate language-production anxiety (Horowitz, 2019; Peterson, 2012b; Reinders & Wattana, 2014). The social gameplay mechanics in MMORPGs have been associated with TL growth in a number of categories including, but not limited to, vocabulary acquisition (Bytheway, 2014; Huang & Yang, 2014; Li, 2019; Peterson, 2010a; Rankin, McNeal, Shute, & Gooch, 2008; Thompson & von Gillern, 2020), speaking skills (Kongmee, Strachan, Pickard, & Montgomery, 2012; Koptur, 2016; Suh, Kim, & Kim, 2010), production confidence (Reinders & Wattana, 2014; Reinders & Wattana, 2015; Sylvén

& Sundqvist, 2012), and improved academic performance in the TL within academic institutions (Sylvén & Sundqvist, 2012).

These linguistic gains in the TL observed in language learners who use socially oriented digital games such as MMORPGs may be interpreted as being a direct result of learning that is taking place from a sociocultural and interactionist perspective, whereby the framework of the MMORPG encourages social exchanges which are beneficial for SLA processes to occur (Chotipaktanasook & Reinders, 2018; Peterson, 2010a; Peterson, 2012b; Peterson, 2016). A specific element that this thesis attempts to address is whether the social gameplay mechanics in MMORPGs are solely responsible for these observed gains in the TL or whether other mechanisms are working in conjunction with these mechanics in order to amplify desirable pedagogical outcomes in SLA contexts. To address this issue, the efficacy of single-player RPGs to elicit SLA is compared directly to MMORPGs.

2.2.2 Non-social, single-player RPGs

Single-player RPGs, in particular those whose content and gameplay mechanics revolve around mission-oriented objectives within a narrative structure, share many fundamental gameplay characteristics with MMORPGs. These overlaps in gameplay content/mechanics that do not rely specifically on social networks may also aid in SLA processes. Potentially beneficial gameplay mechanics for SLA include in-game tasks centered on the completion of missions, whereby scaffolding, corrective feedback, and negotiation of meaning are often used to help progress players through the game (Monfared, Cervantes, Lee, & Jackson, 2018; Rankin & Shute, 2010). Furthermore, contextualization, (Hitosugi, Schmidt, & Hayashi, 2014; Reinhardt, 2017), repetition, (Coxhead & Bytheway, 2015; Heidt, 2020), and multimodal presentation (i.e., audio, visual, and textual presentation) of linguistically authentic in-game dialogue may aid in the long-term retention of TL vocabulary (Bisson, Van

Heuven, Conklin, & Tunney, 2014). From a psycholinguistic perspective, even non-social gameplay mechanics common in many digital RPGs may be beneficial for SLA.

Even when considering the potential benefit of certain elements of non-social gameplay mechanics in single-player RPGs, such digital games may not be completely divorced from the sociocultural/interactionist benefits of gameplay observed in MMORPGs. Indeed, the motivation to use the TL to talk about the game in social circumstances outside the game may be aided by the intrinsic motivation associated with digital games which, in turn, may benefit SLA (Birk, Atkins, Bowey, & Mandryk, 2016; Bostan, 2009; Dickey, 2007; Dörnyei, 1990). Furthermore, the quasi-interactive social exchanges with non-playable characters (NPCs) used in many single-player RPGs may also constitute an artificially constructed social exchange in the TL which may be beneficial to the SLA process (Cornillie, Clarebout, & Desmet, 2012b; Pica, 1987). In order to begin to address the relevance of using single-player RPGs like LIS as the primary medium of instruction in formal foreign language educational contexts, the following section will compare side-by-side how a number of the benefits to SLA observed in MMORPGs compare with the gameplay mechanics of single-player RPGs on the whole and more specifically with the gameplay mechanics of LIS.

2.3 Corrective feedback and negotiation of meaning

Corrective feedback is vital to SLA because it provides the language learner with knowledge regarding the incorrect use or incorrect processing of linguistic forms (i.e., word choice, pronunciation, conjugation, etc.) (Li, 2010). With such knowledge, the language learner is functionally primed to receive the correct version of incorrectly used/processed linguistic forms so that the same mistake is not repeated. The corrective feedback process is achieved either through explicit feedback by another speaker or through implicit feedback such as by self-correction by observation and noticing of the correct linguistic forms used by other speakers (Ellis, Loewen, & Erlam, 2006). As the process of SLA is fraught with misunderstandings, misinterpretations, and general misuse of the syntax and morphology of the TL, a cumulative process of corrective feedback may help push language learners into higher levels of TL proficiency. Furthermore, because corrective feedback can occur both explicitly (e.g., through direct interaction with more capable peers) and implicitly (e.g., through observation of authentic linguistic input), the medium through which corrective feedback may be given is flexible. Circumstances conducive to corrective feedback may therefore include active social exchanges, such as through conversation, or passive TL interaction, such as through reading text. The following section will discuss how the gameplay content/mechanics of MMORPGs, when played in the TL, may be conducive to corrective feedback for language learners.

2.3.1 Corrective feedback in MMORPGs

When a language learner plays an MMORPG on an English language server, the majority of network players will be either native speakers or highly proficient second language (L2) speakers of English (Bryant, 2006). This means that from a sociocultural/interactionist perspective, a language learner playing an MMORPG on an English language server will be primarily interacting with linguistically more capable peers, which, in turn, may encourage SLA growth through such processes as corrective feedback (Peterson, 2012b; Sylvén & Sundqvist, 2012). Due to the fact that many MMORPGs make use of microphones in order to engage with other network players, a language learner playing on an English language server may receive explicit corrective feedback when the mispronunciation of lexical items hinders necessary information exchanges in order to progress in the game. Similarly, when language learners misuse grammatical forms with other network players, whether spoken or through in-game texting features, in such a way that hinders progress in the game, explicit feedback may be given by other players. These suppositional examples of corrective feedback may not necessarily occur with the aim of improving the language learner player's English, but simply for the fact that cooperation is often necessary in order to progress successfully through the game (Pietschmann, Liebold, & Valtin, 2017).

Apart from the explicit corrective feedback that language learners may receive while playing an MMORPG in the TL, another possible route for corrective feedback to take place is through passive observation and intentional self-correction on the part of the language learner. Active linguistic exchanges occur between network players with or without active participation on the part of the learner. Indeed, some language learners who have experienced linguistic gains as a result of playing MMORPGs have identified the specific strategy of passive observation of TL exchanges that occur between other network players (Bytheway, 2014; Li, 2019). Through such observations of authentic TL exchanges between more capable peers, the language learner is provided with the opportunity to recognize the correct use of a linguistic form (e.g., pronunciation, conjugation, word use, etc.) that he/she may have been using incorrectly up to that point, whereby self-correction may take place. These self-corrections may afterwards be re-integrated back into active language production as the player continues to play the game.

This process of implicit corrective feedback also need not necessarily occur as a result of the observation of other network players and may instead occur from the observation of authentic TL use built into the narrative content and gameplay mechanics. In studies analyzing the SLA strategies language learners employ when playing MMORPGs, the strategy of using in-game dialogue with NPCs (i.e., not with other network players as dialogue with computer-controlled NPCs are based on pre-established scripts) and mission briefings to improve their language skills was highly prevalent (Bytheway, 2014; Li, 2019). Furthermore, as is often the case with many modern MMORPG titles, the multimodal presentation of linguistic information through voice-acted content with subtitles and in-game text is widespread (Newon, 2011) and has been associated with linguistic gains in the TL in relevant literature (Boers, Warren, Grimshaw, & Siyanova-Chanturia, 2017; Erdener, 2016).

2.3.2 Corrective feedback in single-player RPGs

The case may then be made that because the process of corrective feedback in MMORPGs is not necessarily dependent on inter-player communication, then other digital games that make use of similar gameplay content/mechanics such as single-player RPGs could arguably achieve a similar process of implicit corrective feedback. One key issue with the claim that single-player RPGs may elicit implicit corrective feedback is that due to the fact that there is no social element to single-player RPGs, and hence no opportunity for a language learner to actively produce self-corrections based within the context of the game, all self-corrected linguistic production in the TL would be relegated to out-of-game contexts (Broussard, 2012). However, the argument may be made that linguistic production in the TL is not the only method to demonstrate the benefits of corrective feedback. By employing a system which requires the language learner to concretely demonstrate comprehension of a previously misunderstood TL linguistic exchange after receiving corrective feedback, linguistic growth in the TL may be clearly demonstrated (Ellis, Loewen, & Erlam, 2006).

Within the non-social gameplay mechanics common to both MMORPGs and single-player RPGs, there is a type of corrective feedback which corrects not for linguistic production but rather for linguistic comprehension. In many story-driven RPGs, certain actions need to be performed in order to move the story forward (e.g., going to a certain location), and instructions on how to perform that action must be comprehended (e.g., receiving information on where a certain location is). When these actions are not completed, the gameplay mechanics may provide a certain level of corrective feedback in the form of hints or more explicit instructions on what the player must do in order to progress in the game. It is precisely these types of gameplay mechanics that constitute a type of corrective feedback for linguistic comprehension in SLA contexts. As corrective feedback mechanics which require language learners to carry out specific tasks related to in-game narrative content do not require TL production, single-player RPGs such as LIS may also be beneficial in SLA contexts, even if limited to eliciting growth in TL comprehension.

2.3.3 Corrective feedback in LIS

When examining a single-player RPG like LIS, there are several prime examples of corrective feedback for linguistic comprehension that are present within the gameplay mechanics. LIS accomplishes this type of corrective feedback for linguistic comprehension through environmental clues, explicit instructions, and gameplay mechanics such as the time reversal mechanics. The time reversal mechanics are unique to LIS and allow the player to reverse time in order to remake in-game dialogue choices. The following sections will provide concrete examples of the environmental clues, explicit instructions, and gameplay mechanics which elict corrective feedback for linguistic comprehension in LIS.

Environmental clues which reinforce explicit in-game instructions are common throughout the game. The first set of linguistic instructions that the player receives at the start of the game involves moving the character Max to a nearby lighthouse. Although the dialogue tells the player exactly what must be done in order to complete this task, for a language learner playing LIS in English, comprehension of the explicit instruction is not a guarantee. However, the in-game environment immediately reinforces these instructions with an environmental clue. Moving Max forward reveals a signpost with the word "lighthouse" accompanied by an arrow pointing in the exact direction the player must go in order to arrive at the lighthouse as well as an image of the lighthouse itself. Such environmental clues functionally serve to correct any misunderstanding of the explicit linguistic instructions (i.e., to move Max to the lighthouse) while simultaneously allowing a language learner player to reassess the meaning of any miscomprehended linguistic forms.

Another process by which LIS elicits corrective feedback for linguistic comprehension is through explicit instructions when the player is no longer progressing through the game. The player is also introduced to this process at the very start of the game. After successfully reaching the lighthouse, Max suddenly wakes up and finds herself at school in the middle of a lecture in photography class. The teacher, Mr. Jefferson, begins to talk about the historical origins of photographic "selfies." This lecture point acts as a linguistic clue indicating that the player must take a selfie with the Polaroid camera on the desk in order to move the story forward. Should the linguistic clue go unnoticed, Max provides the player with a more specific prompt with the inner dialogue telling the player to take a selfie. Should the added linguistic instruction from Max telling the player to take a selfie still go unnoticed, the game then provides explicit instructions on how to use in-game controls to take a selfie. Just as with the environmental clues, the gameplay mechanics in LIS provide explicit corrective feedback when the player has failed to comprehend other linguistic commands in the game, which in turn allows a language learner player to reassess the meaning of any miscomprehended linguistic forms that impeded progress through the game.

Finally, the time reversal mechanics allow the player to rewind short narrative sequences in order to redo specific in-game actions or to provide different script choices to in-game dialogue. LIS is a choose-your-own-adventure RPG, a genre which relies on in-game choices, the sequencing of which will produce one of many possible endings or narrative paths (Tyndale & Ramsoomair, 2016). The choice-based gameplay mechanics also function to facilitate in-game dialogue puzzles which the player must solve to continue playing. After the player has taken a selfie at the start of the game, Mr. Jefferson asks Max to join the class discussion on the topic of selfies by asking her to name the process that gave birth to the first photographic self-portraits. Max does not know the answer to the question, and an NPC named Victoria provides the correct answer to the question. The correct information can then be used by reversing time and then choosing the correct answer to the question. The time reversal mechanics functionally provide the player with the ability to reassess the linguistic information provided by the game and then select the correct script choices to mandatory in-game dialogue puzzles. Therefore, a language learner player may reassess the meaning of linguistic forms through this process of corrective feedback for linguistic comprehension.

2.3.4 Summary of corrective feedback

In MMORPGs, the social gameplay mechanics create the impetus for explicit corrective feedback to occur between language learners and linguistically more capable peers because players need to have comprehensible linguistic communication with one another in order to progress in the game. Self-correction may occur implicitly through the observation of other inter-player communication or game-specific dialogue conducted with NPCs. Single-player RPGs also provide a range of multimodal linguistic input which may provide language learners with the opportunity to self-correct misused linguistic forms outside the context of the game. Although single-player RPGs do not elicit in-game TL production, they may provide corrective feedback for TL comprehension. In LIS, this process may be accomplished with gameplay mechanics that allow the language learner to reassess the meaning of miscomprehended linguistic input from the game through environmental clues, explicit instruction, and the ability to remake in-game script choices. In SLA contexts, such gameplay mechanics will arguably lead to linguistic growth even in the absence of social gameplay mechanics which allow for in-game TL production.

2.3.5 Negotiation of meaning

Another process conducive to SLA that has been identified within the genre of MMORPGs are gameplay mechanics which may elicit negotiation of meaning (Dixon, 2104; Peterson, 2012a, Peterson, 2016). Negotiation of meaning is a process that a language learner actively engages in to arrive at a clearer understanding of said linguistic forms in the TL. The process of negotiation may occur through instances of clarification (e.g., asking for/giving more specific information), rephrasing (e.g., swapping out technical language for standard vocabulary), repetition (e.g., repeating the necessary steps to complete a task), and confirmation (e.g., double checking to make sure that a command was understood correctly) (Fernández-García & Martínez-Arbelaiz, 2002). As the process of negotiation of meaning often requires some kind of linguistic interaction, it is theorized to be beneficial to SLA from an interactionist perspective (Chapelle, 2006; Fuente, 2003; Peterson, 2010b). MMORPGs and single-player RPGs whose gameplay mechanics elicit negotiation of meaning may therefore also elicit linguistic growth when played in the TL from the same theoretical standpoint.

2.3.6 Negotiation of meaning in social and non-social RPGs

In MMORPGs, negotiation of meaning is elicited through the social gameplay mechanics. Just as corrective feedback in MMORPGs is driven by the need for comprehensible linguistic cooperation in order to progress through the game, negotiation of meaning in MMORPGs may occur when groups of network players have to clarify mission objectives, rephrase in-game technical language or regional colloquialisms, repeat the steps necessary to accomplish a mission, or clarify the effect of a specific in-game item. In SLA contexts, when a language learner engages with MMORPGs in the TL, these processes of negotiation of meaning that are routinely carried out between groups of network players in the TL have been observed to be conducive to SLA (Rankin, Gold, & Gooch, 2006; Peterson, 2012a). However, just as with the process of corrective feedback in MMORPGs, the RPG gameplay mechanics that operate in tandem with the social elements in MMORPGs may also work to elicit negotiation of meaning. These non-social negotiation processes may occur between the player and the artificially constructed social environment, allowing for linguistic interactions to take place without the need for other human players.

The linguistic exchanges of information that occur during negotiation of meaning need not necessitate the active production of language and may include the use of reference material as a medium of negotiation (Cornillie, Jacques, De Wannemacker, Paulussen, & Desmet, 2010). Operating under this definition, any item which contains information that can be used for linguistic cross referencing (e.g., textbooks, operation manuals, etc.) could potentially elicit TL gains through negotiation of meaning. Returning to the case of the non-social gameplay mechanics of MMORPGs, the in-game environment clearly elicits non-social linguistic interactions with NPCs, game menus, and narrative structure whereby negotiation of meaning may occur. When the scope of the gameplay mechanics conducive to negotiation of meaning within MMORPGs is stretched to include even the non-social elements, then single-player RPGs which utilize similar gameplay mechanics may also elicit similar benefits for SLA.

Identifying how negotiation of meaning as defined above (i.e., clarification, rephrasing, repetition, and confirmation) functions in digital games without social interactions will help to establish the validity of using single-player RPGs in SLA contexts. Just as with the non-social gameplay elements in MMORPGs, single-player RPGs use in-game environments to elicit interactions. Such in-game interactions (e.g., with NPCs, game menus, narrative structure) often require the player to understand linguistic information in order for a successful interaction to take place. For a language learner playing in the TL, the linguistic information provided by such gameplay mechanics as automated script responses from NPCs may be used to negotiate the meaning of unknown TL linguistic forms. The following section will detail how the gameplay mechanics of LIS function to create just such an environment conducive to negotiation of meaning.

2.3.7 Negotiation of meaning in LIS

When playing a TL digital game that requires the player to complete specific in-game objectives, the miscomprehension of TL linguistic forms can impede successful progression through the game. Choose-your-own-adventure, single-player RPGs such as LIS incorporate gameplay mechanics which provide the player with the opportunity to negotiate the meaning of misunderstood TL linguistic forms. Gameplay mechanics in LIS which elicit negotiation of meaning are evidenced in the following section with a specific example of an in-game objective whereby Max must retrieve her friend Warren's flash drive from her dormitory room and return it to him. The analysis of how this process of negotiation of meaning may function during gameplay relies on the hypothetical supposition that a language learner does not know the meaning of the word "flash drive," thus hindering the player's ability to complete the in-game objective.

The in-game objective starts with the notification that Max has received a text message from her friend Warren. Should the player not notice or not know how to proceed, the game provides explicit corrective feedback telling the player that he/she needs to read the text message from Warren in order to progress through the game. A series of text message exchanges then occur between Max and Warren which outline the in-game objective of returning Warren's "flash drive." In line with the supposition for analysis, the language learner does not know what the word "flash drive" means, but linguistic clues as to what the item is used for are provided in the text message exchanges. Warren states that he needs his flash drive back because he needs "info" and "space." Should the player scroll through the history of text message exchanges between Max and Warren, another linguistic clue is provided as to what the item "flash drive" is used for when Warren encourages Max to check out something titled "NEKromantik" which is on the flash drive. Assuming the language learner in this hypothetical scenario understands the terms "info" and "space," the assumption may be made that the unknown word "flash drive" is something which can store (inferred from the word "space") files (inferred from "NEKromantik" being *on* the flash drive).

After the player is finished reading through the text message exchanges between Max and Warren, Max provides the player with information as to where the "flash drive" may be obtained by stating that she needs to go to her dormitory. Upon entering Max's room, the player is confronted with a room full of interactive objects. Assuming the language learner still does not know precisely what the word "flash drive" means but is equipped with the knowledge that the item is somewhere within Max's room, then the player may engage in a search for the item by interacting with the items in Max's dormitory room. Assuming that the language learner has made the connection provided by the contextual clues in the text message exchanges between Max and Warren that a "flash drive" is an item which can be used to store computer files, then the player may start their search at the laptop sitting upon the desk. Even if the language learner has not pieced together what the item "flash drive" is used for, a diligent search through the myriad of interactive items in Max's dormitory room will eventually uncover the next step in the in-game objective by arriving at the laptop. In front of Max's laptop, there is a pink post-it note from an NPC named Dana who borrowed the flash drive to watch some "flix" while studying. Max then directs the player to continue the search for the flash drive in Dana's room. The search for the flash drive in Max's dormitory room provides further environmental and linguistic clues as to what the item is used for. Not only is the post-it note placed in front of the laptop, providing further evidence that the "flash drive" item is used to store computer files, but the post-it note itself states that

Dana used the flash drive to watch “flix,” another word the language learner may be unfamiliar with but can assume that a screen, such as would be provided with a laptop, would be necessary in order to physically watch the file.

When Max finds Dana’s dormitory room, the search for the flash drive begins anew. Assuming that the language learner has been able to put together the multitude of environmental and linguistic clues as described above to arrive at an assumption that a flash drive is a storage device for computer files, the player may go straight to the laptop on Dana’s desk and retrieve the flash drive. Just as with the search in Max’s room, should the language learner still be unsure as to the function of the flash drive, the player would have to search through the interactive items in Dana’s dormitory room before eventually coming across the flash drive on the desk. Upon picking the flash drive up, an icon of a flash drive along with the acronym USB appears in the upper right-hand corner of the screen. With a visualization of the flash drive along with the added information that a flash drive is a USB device that interfaces with computers and stores files, it should be abundantly clear as to the meaning of “flash drive” to any language learner playing LIS in English. Conversing with Dana even provides the language learner with even more linguistic clues as to the use of the flash drive when Dana states to Max that many “films” on Warren’s flash drive were obscure and that she had not even heard of them.

2.3.8 Summary of negotiation of meaning

The hypothetical scenario described above of a language learner playing LIS in English demonstrates how the gameplay mechanics allow the player to negotiate the meaning of unknown linguistic forms, in this circumstance the word “flash drive.” The gameplay mechanics may function, as in the scenario described above, to allow the language learner to negotiate unknown TL vocabulary by eliciting instances of clarification, rephrasing, repetition, and confirmation. For clarification, the player has unlimited access to Max’s journal (i.e., the in-game menu) as well as the text message exchanges between Max and Warren to clarify the mission objectives. For rephrasing, the word "flash drive" is referred to in a number of different ways in terms of both function and title. Warren's flash drive is rephrased as "USB," and the function of the flash drive is rephrased with the words "info/space," "flix/watch," and "files/films." For repetition, both the word and in-game objective are repeated numerous times until the flash drive is successfully retrieved and returned to Warren. These repetitions occur as in-game menu reminders of mission objectives and in-game dialogue with NPCs. Finally, for confirmation, the player is given a number of environmental and linguistic clues as to what the word "flash drive" means. Assuming that the language learner is not employing a brute-force strategy of checking every single item on every single screen in order to progress the storyline forward, the player may utilize these environmental and linguistic clues in order to arrive at an assumption of what the item is and where it is located. Even without social gameplay mechanics, single-player RPGs such as LIS display key characteristics which allow language learners to negotiate the meaning of unknown TL linguistic forms. The incorporation of such gameplay mechanics would also in turn provide evidence for the pedagogical viability of using single-player RPGs in SLA contexts.

2.4 Contextualization of language

Another process that has been identified as being integral to SLA is the contextualization of language (Godwin-Jones, 2018; Moore, 2018). When TL linguistic forms are learned out of context such as through rote memorization for purposes divorced from authentic linguistic applications (e.g., memorizing TL linguistic forms for the purpose of taking a test), language learners may fail to make relevant connections to how TL linguistic forms are used in socially authentic contexts (e.g., within a profession). In a general sense, this means that learning the TL without contextualization may limit social interactivity in the TL and lead to the use of avoidance strategies (MacIntyre, Clément, Dörnyei, & Noels, 1998; Piechurska-Kuciel, 2011; Yashima, 2002). A lack of contextualization in the SLA process may also have more specific consequences on the ability for language learners to develop the necessary skills to infer the meaning of unknown TL linguistic forms in authentic social settings (Pearl & Goldwater, 2011) as well as on the ability to mobilize relevant knowledge of TL linguistic forms as needed (Godwin-Jones, 2018). As language is primarily a tool used to navigate social structures, many researchers in the field of applied linguistics argue for the need to construct in-class lessons that incorporate authentic contextualization of the TL (Fernandes, Leite, Mouraz, & Figueiredo, 2013). The following sections will introduce how contextualization of language is used in SLA pedagogical methodologies, then use the concepts highlighted therein to frame a discussion about how these processes might be at play in digital games.

2.4.1 Task-based language learning

Although there is no one-size-fits-all approach to contextualize SLA pedagogy, task-based language teaching (TBLT) has been advocated as a teaching methodology that may effectively introduce a level of contextualization into the language learning process in formal educational environments (Long, 2016). TBLT is a methodology that focuses on the use of authentic language, derived in principle from the employment of simulations of authentic social interactions in the TL (e.g., going to the doctor). Within the methodological context, the language learner completes meaningful tasks in the TL (e.g., making an appointment on the phone, filling out a patient history application, and communicating with a doctor about symptoms/treatment options). The TBLT curriculum therefore employs a pedagogical goal and requires students to accomplish a series of tasks using the TL to arrive at that goal. In order to prepare students to be able to accomplish these tasks in the TL, any number of teaching materials may be employed (e.g., vocabulary/grammar lists of TL terminology relevant to the task) whereby all materials taught in class relate to a contextualized system of progressing events to reach the pedagogical goal.

In addition to the socially authentic application of the pedagogical goal, both the type of tasks and method of administration are vital to ensure desirable pedagogical outcomes. According to Rod Ellis (2003), a task should display four general qualities: 1) a task should focus on practical applications, 2) a task should incorporate an information gap whereby the students must utilize practical linguistic skills to fill in such gaps, 3) the students should have control over which linguistic skills to use in order to successfully complete the task, and 4) the goal of the task is not specifically centered on linguistic gains. TBLT curricula may also function to elicit SLA by correlating individual tasks with one another to build interlinking skills that ultimately allow the students to accomplish the overarching pedagogical goal (Van Avermaet, Colpin, Van Gorp, Bogaert, & Van den Branden, 2006). Tasks may therefore be thought of as subsets of the overall goal of the curriculum, the sum of which culminate in the students being better able to accomplish the pedagogical goal within an authentic context. From an SLA perspective, such a system of interlinking tasks working in concert to create a larger skill may be particularly conducive to language growth because it contextualizes the wide variety of linguistic forms employed in each of the tasks and encourages comprehension of each form used before being able to move on to the next task, whereupon the skills utilized in previous tasks may be reused and elaborated upon in different contexts (Nagle & Sanders, 1986). Many digital games, including MMORPGs and single-player RPGs, also incorporate gameplay mechanics which mirror TBLT as defined above (Sykes, 2014; Vegel, 2018).

2.4.2 Contextualization of language in social and non-social RPGs

Most well-designed digital games, as argued by Paul Gee (2003a), function as highly efficient learning systems in which the player undergoes a series of tasks of increasing difficulty designed to arrive at the pedagogical goal of the player being able to play the game well. From an SLA standpoint, gameplay mechanics that require the use and/or comprehension of linguistic elements to complete such tasks that increase in complexity as the game progresses would closely resemble the concept of scaffolding, a pedagogical method argued to be conducive to SLA (De Guerrero & Villamil, 2000). The concept of scaffolding, as it relates to digital games, involves the completion of in-game tasks which increase with difficulty as the player progresses through the game, often requiring the player to utilize skills acquired from previously completed in-game tasks.

Using the four characteristics of what an effective task within the context of a TLBT curriculum entails as defined by Rod Ellis (2003), a digital game may help contextualize SLA by 1) providing a practical application of the TL by way of providing an in-game task with clearly defined goals, 2) encouraging players to utilize linguistic information/communication to solve in-game information gap problems, 3) giving the players agency over how they solve in-game tasks, and 4) prioritizing the completion of the in-game task over the improvement of general TL skills. Another potential benefit of utilizing digital games to contextualize SLA is that, in contrast to TBLT curricula in which the completion of each task is not strictly controlled to ensure that each student has comprehended the task before moving on to the next one, the gameplay mechanics of many digital games do not allow the player to progress until each task is completed, which, from a theoretical standpoint, would encourage better pedagogical outcomes.

With social gameplay mechanics that necessitate the cooperation between network players to complete in-game tasks, MMORPGs may be highly conducive to SLA in the way that they contextualize language for use as a practical in-game skill in order to progress through a series of increasingly difficult in-game tasks (Reinhardt, 2017; Thomas, 2012). One example of how this contextualization process may function in an MMORPG is through the application of in-game resources to solve specific problems which hinder the player's progress through the game. Given a hypothetical scenario in which the player must retrieve an item from a specific dungeon, one obstacle that the gameplay mechanics may put in place is to populate the dungeon with a specific type of enemy. When the player has just started to play the game and the player's character is at a low level, certain enemies may present a formidable obstacle to completing the in-game task. This hypothetical dungeon populated with formidable enemies may be thought of as a type of information gap and may be solved in a number of ways conducive to SLA processes. Some solutions may involve finding specific information (e.g., enemy weaknesses) or eliciting cooperation from other players which would allow the low-level player to complete the dungeon. Language learners who are playing the game in the TL are therefore effectively undergoing a TBLT learning process by using unspecified linguistic forms to accomplish a practical in-game task (i.e., completing a dungeon). The linguistic skills acquired by such in-game tasks may then be utilized in increasingly complex scenarios as he/she progresses through the game.

The contextualization of language engendered by MMORPGs is not limited to social gameplay mechanics. Indeed, many single-player RPGs arrive at similar levels of linguistic contextualization through the employment of in-game interactions with NPCs. Similarly, given the same hypothetical scenario described above of a low-level player engaging with a difficult dungeon, a language learner player may employ similar information search strategies in single-player RPGs to defeat the dungeon. The key difference is that language learner players cannot elicit direct in-game cooperation from other human players. However, the TBLT learning process of using unspecified linguistic forms to accomplish a practical in-game task remain largely the same with or without social gameplay mechanics. The following section will outline a specific example of how the gameplay mechanics of LIS contextualize language use in a manner conducive to TBLT learning.

2.4.3 Contextualization of language in LIS

The gameplay mechanics of LIS require players to utilize in-game information to successfully navigate through social exchanges. These social exchanges require the successful selection of script choices during interactions with NPCs. As the game progresses, these social exchanges become more and more complex, but through the process of scaffolding, the player may become more attuned not only to what in-game information they need to pay attention to but also to which script choices need to be made in order to successfully progress through the game. Language may be contextualized in single-player RPGs by way of providing the player with linguistic information necessary to progress through the game which results in the player taking a non-linguistic in-game action (e.g., navigating to a specific location). With the script choice mechanics in LIS however, the player experiences first-hand examples of how social exchanges in the TL may play out depending on the input given. Moreover, the player is given the option to actively alter the trajectory of these social exchanges through the time reversal mechanics.

Within the in-game objective of retrieving Warren's flash drive from Max's dormitory room described in the previous section, there is an obstacle which functions as an information gap task, whereby the player must engage in a script choice exchange with multiple outcomes. After discovering that Dana borrowed Warren's flash drive from Max's room, Max must go and retrieve the flash drive from Dana's dormitory room. Upon entering the hallway, Dana's roommate Juliet is shown locking Dana inside of the room due to an argument that has arisen between the two of them, creating an obstacle for Max to retrieve the flash drive. The player must therefore help solve the argument between Juliet and Dana, the first step of which is to engage Juliet and find out what the problem between the two of them is. Probing the situation with the script choice mechanics reveal that Juliet is under the impression that her boyfriend Zach has been "sexting" (i.e., texting with overt sexual context) with Dana. The player also has the choice of remaking script choices with the time reversal mechanics to uncover potentially useful information hidden behind different script choices.

After the initial social exchange between Max and Juliet, regardless of the conversation path, Max will continue to probe for more information which may help her solve the dilemma by asking Juliet how she found out about her boyfriend's supposed impropriety with Dana. However, Juliet questions Max's motives by claiming that Max does not care about anyone else and tests Max by asking her if she knows her last name. At this point, the player must make a choice between four options of last names. After incorrectly guessing Juliet's last name, Juliet refuses to talk to Max anymore. In order to proceed any further in the game, the player must understand from the social exchange that Juliet is upset that Max did not know her last name and that in order to avoid upsetting Juliet, the conversation must be redone with the time reversal mechanics. If this intricate nuance within the social exchange is not noticed, as might be the case with a language learner who has not fully understood the complexities of the conversation, the player will be unable to progress any further in the game. The gameplay mechanics still do, however, continue to provide corrective feedback to keep players progressing through the game through explicit instructions stating that information from the social exchange can be used to alter the trajectory of the conversation.

Upon figuring out that the conversation must be redone from the beginning with the aid of the time reversal mechanics, the player must then utilize the content of the previous linguistic exchanges (assuming that the player did not happen to guess the correct name on the first try) and prove to Juliet that she knows her name, whereupon Juliet becomes less suspicious of Max. Juliet then provides the player with clues as to the source of the argument by telling Max that the character Victoria told Juliet about Zach's impropriety. As has been established so far in the character development of Victoria up to this point in the game, Victoria has been positioned to be heavily antagonistic to Max. Max is therefore suspicious that Victoria has an ulterior motive and provides the player with linguistic clues that the key to solving the information gap puzzle has something to do with Victoria. Even more linguistically explicit instructions come shortly thereafter from Dana who shouts to Max that Victoria is behind this misunderstanding and that the proof is in Victoria's dormitory room.

The next step that the player has to take in order to solve the information gap puzzle is clearly laid out both linguistically and contextually. Should a language learner playing the game in English be unable to pick up upon these linguistic and contextual instructions on how to progress through the game, the player can rely on the time reversal mechanics in order to engage in the social exchanges with both Juliet and Dana again and again, as many times as needed until the next step of the puzzle becomes clear. Walking into Victoria's room, the player should have a strong understanding of what needs to be found, namely proof that Victoria lied to Juliet when she told her that Juliet's boyfriend Zach was engaging in lewd text message exchanges with Dana. Equipped with such contextual knowledge, a logical point to start the search amongst the myriad of interactive objects within the room is Victoria's laptop, as it may contain written evidence of Victoria's subterfuge. Even if a language learner has failed to understand enough of the linguistic nuance in the conversation between Max and Juliet to narrow down the scope of the search, as long as the general objective of looking for something in Victoria's room has been understood, the correct item will eventually be stumbled upon. Upon searching Victoria's computer, Max finds an email that Victoria has written to a character named Taylor in which she describes how she purposefully lied to Juliet about Dana engaging in lewd text message exchanges with Zach. She further goes on to describe her motive behind straining Juliet and Dana's relationship, revealing that she and Zach once dated and that she feels as if Juliet stole Zach away from her.

Within the context of the in-game information gap task, the player should understand that this email is what is needed to prove to Juliet that Dana is innocent, which will in turn solve the conflict between the two roommates and allow Max to enter the room to retrieve the flash drive. Furthermore, upon discovering the email, Max not only states that this is the email that she needs to show to Juliet, but she also explains how the player is supposed to relay this information to Juliet by stating that she needs to print the email and physically hand it over to Juliet. Therefore, simply discovering the email does not automatically solve the obstacle, as the player must understand both the context of the situation as well as the linguistic clues provided by Max. The player must exit out of Victoria's web browser and then interact with Victoria's laptop one more time, whereupon a new action of "print e-mail" appears. In other words, simply stumbling upon the necessary item is not enough, and the player must comprehend the linguistic clues to carry out specific actions within the context of the task in order to progress through the game. Only after printing the email and handing it over to Juliet is the argument between the two roommates resolved, whereupon Max may finally enter the room in order to retrieve Warren's flash drive and complete that in-game objective.

2.4.4 Summary of contextualization of language

The degree of control that a player has over the course of social exchanges within the gameplay mechanics of LIS combined with the utilization of these social exchanges as information gap tasks within the context of in-game objectives provides a framework similar in structure to TBLT. Furthermore, the scaffolding of complexity as the player moves through the game encourages improvement in how proficiently the player interacts with these information gap social exchanges. In particular, it is the information gleaned from social exchanges that can help solve information gap puzzles which may work to incentivize both engaging in these types of interactive social exchanges and using the time reversal mechanics to try and uncover new information from previously unselected conversation paths for practical in-game purposes. Therefore, when played in the TL, single-player RPGs with gameplay mechanics similar to the interactive script choice mechanics in LIS may help to contextualize processes conducive to SLA.

2.5 Authentic linguistic input

Alongside the argument calling for the authentic contextualization of SLA pedagogy runs the parallel argument calling for the use of authentic linguistic input. The term "authenticity" as it applies to SLA pedagogy may be interpreted as the authenticity of language, authenticity of task, and authenticity of situation (Breen, 1985; Taylor, 1994). Just as TBLT calls for tasks whose primary focus is not on linguistic outcomes but rather on realistic scenarios in which natural language use arises, authenticity of linguistic input in SLA pedagogy may be defined as the employment of materials not principally designed for SLA and which exemplify how the TL functions for real-life communicative purposes (Nunan, 1988). Principally such materials would include those created for use by native speakers and highly proficient L2 speakers in the TL culture (e.g., films, novels, newspapers). However, some researchers have called for a broader definition of authenticity to include materials with more specific SLA applications such as graded readers (Chavez, 1998; Widdowson, 1998). The concept of linguistic authenticity in SLA contexts is therefore elastic. However, for the purpose of analyzing the potential authenticity of utilizing digital games in formal educational environments, the following section will utilize a broad definition of the concept to include any material which exposes language learners to practical linguistic input on how native speakers use the TL for everyday communication. In particular, the following discussion will delve into how MMORPGs and single-player RPGs like LIS may function to expose students to high levels of authentic linguistic input in the form of nonstandard language use, namely slang, texting shorthand, and cultural referencing.

2.5.1 Linguistic authenticity in SLA contexts

Considering the prevalence of nonstandard language use in everyday communication, slang, texting shorthand, and cultural referencing certainly all fall under the category of authentic language. It is therefore even more surprising that more time is not devoted to nonstandard language in SLA contexts despite calls for increased attention to such underrepresented linguistic forms (Dinçay, 2012). Indeed, the moniker "nonstandard" itself erects a barrier to the inclusion of slang, texting shorthand, and subcultural referencing into "standardized" SLA curricula. However, the title of "nonstandard" does not stop native speakers from using such speech patterns for communicating "standard" information. This is not to say that SLA curricula should primarily be centered on nonstandard language instruction, however pedagogical methodologies that focus entirely on standard language use for not so practical purposes such as standardized test taking do not adequately prepare students for interaction with native speakers in communicative contexts as well as with the TL culture. One way to ensure that language learners are prepared for the communicative realities within the TL culture is through the employment of materials produced for use primarily in the TL culture.

Due to the breadth of authentic materials present in the TL culture that may be employed for SLA purposes, careful consideration of the pedagogical goals of the curriculum must be considered before selecting an appropriate medium of instruction. Depending on the medium of instruction, linguistically authentic materials may help language learners enhance lexical depth (Gilmore, 2007), improve pronunciation/listening (Gilakjani, 2016), increase intercultural competency (Maiier, 2017), and supplement communication strategies in the TL (Gilmore, 2011). Relevant literature has provided a multitude of studies examining how linguistically authentic media for SLA contexts includes internet message boards (Coppens, Rico, & Agudo, 2013; Nah, White, & Sussex, 2008), film/televised media (Gilmore, 2011; Grode & Stacy, 2015), smartphone applications (Rosell-Aguilar, 2017), graphic novels (Hüseyin & Efecioğlu, 2015; Wrobetz, 2018), and digital games (Peterson, White, Mirzaei, & Wang, 2020; Sykes & Dubreil, 2019; Thorne & Watters, 2013). Of course, trends in pedagogical methods do not necessarily translate into desired pedagogical outcomes, and there are certainly distinct drawbacks to using media designed primarily for native speakers in SLA contexts.

In contrast to instructional materials that are specifically designed for use in foreign language classrooms, the linguistic content of media produced primarily for native speakers will most likely only be comprehensible to relatively advanced language learners. Furthermore, as such TL media were not specifically designed for pedagogical applications, most will not feature the complexity scaffolding present in many well-designed textbooks whereby incremental increases in difficulty move the learner through interlinking lessons which build upon the content of the preceding chapters. In addressing these potential drawbacks, it is again imperative to carefully consider the type of linguistically authentic TL media and the educational environment in which it will be employed. For classes in which the linguistic level of the enrolled students is not high enough to meaningfully interact with language-only media (e.g., newspaper articles), then employing multimodal media that employ visualizations of content such as digital games may help to ease the comprehensibility gap posed by relatively low levels of TL proficiency. Although certainly not the only linguistically authentic media with SLA applications for low-level classes (films/televised media and graphic novels also possess multimodal capacities) digital games also allow the player to interact with the environment which may help to further bridge these comprehensibility gaps. Moreover, most digital games make use of gameplay mechanics conducive to learning such as complexity scaffolding, making digital game genres already identified as being conducive to SLA processes such as MMORPGs and single-player RPGs all the more appealing to be used as sources of authentic linguistic input in language classrooms with varying levels of ability in the TL.

2.5.2 Authentic linguistic input in MMORPGs

MMORPGs provide rich sources of authentic linguistic input with specific regard to nonstandard language use which may come in the form of slang, texting shorthand, or cultural referencing. The social gameplay mechanics which enable inter-player communication through microphones and chat boxes functionally simulate real-life contexts in which spoken and written communication occurs (Dixon & Christison, 2021; Peterson, 2016; Peterson, White, Mirzaei, & Wang, 2020). When a language learner engages with the social gameplay mechanics on a TL server with native speakers of the TL, not only does the player receive ample passive exposure to authentic spoken and written communication in the TL but the player may in turn reproduce authentic language in both in-game and out-of-game contexts (Rankin, McNeal, Shute, & Gooch, 2008; Sylvén & Sundqvist, 2012). Relevant literature identifies six categories of nonstandard language use: loans, shortenings, shifts, blends, composites, and new creations (Algeo, 1999; Ensslin, 2012). Research has also identified regular use of these forms of nonstandard language, a term sometimes dubbed "gamer-speak," by MMORPG players during gameplay (Strong, 2018). These forms of nonstandard language/slang are widespread in everyday communication (Mattiello, 2005) and may therefore represent one avenue to build communication skills in the TL in conjunction with the study of standard TL forms.

Apart from the use of in-game slang in MMORPGs, another form of nonstandard language that MMORPG players may engage in is texting shorthand. Texting shorthand is often used for written linguistic exchanges with the chat box features in MMORPGs. Common examples of English texting shorthand include, "LOL (laugh out loud)," "TTYL (talk to you later)," "LMAO (laugh my ass off)," "IMO (in my opinion)," and, "BRB (be right back)." Having originally evolved through computer-mediated instant messaging technology, texting shorthand has gradually worked its way into everyday use, especially with the spread of mobile text messaging (Tagliamonte & Denis, 2008). Therefore, media that make use of texting shorthand and other forms of nonstandard language described above may be considered valuable sources of authentic linguistic input (Ling, 2005; Plester, Wood, & Joshi, 2009).

Cultural referencing is another unique type of nonstandard language use that MMORPG players will engage in during gameplay which involves linguistic references to a shared cultural heritage which will influence or otherwise replace standard linguistic terminology. The adjectives "leet (sometimes spelled l33t or 1337)" and "noob (sometimes spelled n00b)" are examples of how cultural referencing may function as nonstandard language use within the social sphere of online digital gaming. The term "leet" is a shortening of the term "elite" and the term "noob" is a shortening of the term "newbie," itself a composite of the term "inexperienced newcomer" (h2g2, 2002). Both terms stem from "leetspeak," a unique linguistic system of substitutional orthography using homoglyphs (i.e., two or more glyphs which closely resemble one another) which evolved on bulletin board systems (BBS), an early precursor to the world wide web in use during the 1980s (Nishimura, 2015). Groups of hackers such as the Cult of the Dead Cow are associated with the origin of leetspeak, and it is theorized that the pseudo cryptographic nature of the orthography evolved to enable open discussion of legally questionable topics such as hacking on BBS message boards (Flamand, 2008). The culture of leetspeak eventually spread to the social sphere of digital gaming and into common nonstandard English terminology, whereby the term leet is now synonymous with "a highly skilled gamer," and the term noob is now synonymous with "an unskilled beginner, especially in gaming" (Sjöblom & Aronsson, 2012). Moreover, the use of such nonstandard terminology, especially with the more obscure spellings of l33t/1337 and n00b is an act of cultural referencing which signifies an association with the broader computer culture (Flamand, 2008; Nichol & Blashki, 2008). As leetspeak-derived terminology such as leet and noob have entered the mainstream of nonstandard English use, media which regularly make use of such terminology such as MMORPGs may be able to provide a rich source of authentic linguistic input for language learner players.

2.5.3 Authentic linguistic input in single-player RPGs

Whereas digital games that make use of social gameplay mechanics such as MMORPGs may expose language learners to authentic linguistic input by way of spoken and/or written interactions with native speakers, single-player RPGs cannot provide language learners with access to such direct in-game experiences. The authenticity of the linguistic input in single-player RPGs must therefore be judged solely on the authenticity of the game's linguistic content. When single-player RPGs are played in the language in which they were originally produced in, then the authenticity of the linguistic content is more or less guaranteed. However, when single-player RPGs are played in a language other than the original version, then the authenticity of the linguistic content depends entirely on the quality of the translations. With the rise of and total dominance of Japanese console systems (e.g., NES, SNES, Sega Genesis, PlayStation) from the mid 1980s all the way up to the release of the Microsoft Xbox on November 15th, 2001 (Gurwin, 2021), so too came the rise of digital game titles developed by Japanese gaming companies (Marchand & Hennig-Thurau, 2013). However, Japanese game developers have produced many examples of questionable English translations.

The dominance of Japanese-produced single-player RPGs throughout the late 20th century makes it impossible to ignore the subject of translation quality when analyzing the linguistic authenticity of single-player RPGs in EFL contexts. Perhaps the most famous

English mistranslation from a Japanese-produced game, the single line "All your base are belong to us" perfectly encompasses how translation quality can completely invalidate the pedagogical potential of digital games from an SLA perspective. The world-famous mistranslation is from *Zero Wing* (Toaplan, 1989), a side-scrolling shooter arcade game developed by Toaplan and published by Namco and Williams Electronics in 1990 in North America. As was common in the early days of exporting Japanese digital games to the international market, the translation of *Zero Wing* was handled in-house by non-native English speakers employed by the game's developer Toaplan (Tane & Yamamoto, 2012). The strategy of utilizing in-house employees who were not native English speakers was prevalent even among well-established game developers such as Square Enix, a practice which produced an equally infamous English mistranslation of, "You spoony bard!" from the North American release of *Final Fantasy IV* (Square, 1991). These examples of English mistranslations in Japanese single-player RPGs such as the wildly inaccurate grammatical mistakes listed above are not limited to the early days of gaming market internationalization.

As the commercialization of console-based digital games gained momentum throughout the 1990s with the critical market success of a number of key console systems (e.g., SNES, Sega Genesis, PlayStation, Nintendo 64), the budget allotted to the development of individual digital game titles began to rise in turn (Theodore, 2016). However, bigger budgets and more attention to the importance of international markets still could not immunize Japanese-produced single-player RPGs from English translation gaffes. Arguably the most influential single-player RPG of its era, *Final Fantasy VII* (Square, 1997) had one of the largest development budgets for a single digital game title in the 1990s at approximately $40 million (USD) (Leone, 2017). A number of English translation errors still managed to plague the game's North American release, despite going on to achieve wide-spread critical acclaim. Following the critical success of *Final Fantasy VII*, Square Enix reportedly came to the decision to allocate a higher budget to translations with specific regard to pay more attention to localization (Honeywood, 2007; Leone, 2017). Even with the improvements that larger budgets and larger audiences have brought to the quality of translation in Japanese-produced single-player RPGs, localization errors still run rampant even in more modern titles (Klepek, 2017). Translation quality must therefore be carefully scrutinized when employing internationally produced single-player RPGs in SLA contexts.

While many mistranslations are not severe enough to make the linguistic content of digital games incomprehensible, they do bring into question the pedagogical applicability of translated single-player RPGs in SLA contexts. For that reason, the authenticity of the English used in LIS was thoroughly examined in order to determine its pedagogical applicability in EFL contexts. Although published by the Japanese game developer Square Enix, LIS was developed by the French company Dontnod Entertainment. The script was originally written in French by Jean-Luc Cano and then translated into English by Christian Divine and Cano (Phillips, 2016). Since the gameplay mechanics revolve heavily around in-game dialogue, a large portion of the budget was dedicated to writing and voice acting which produced a highly authentic linguistic rendering of the English language in a single-player RPG (Phillips, 2015). Moreover, the linguistic content of LIS features highly authentic nonstandard English similar to what a language learner may be exposed to in MMORPGs. The following section will detail specific examples of authentic nonstandard English use from LIS.

2.5.4 Nonstandard language use in LIS

Set at Blackwell Academy, a high school in the fictional city of Arcadia Bay in the Pacific northwest of the United States, many of the characters depicted in LIS are teenagers and thus, in line with their character depictions, frequently use authentic nonstandard English terminology. Examples from all five of the six categories of nonstandard language use as described above can be found in abundance throughout the game. On numerous occasions throughout the game, the character Chloe constructs pejorative composites in reference to her stepfather David by affixing a number of derogatory words to “step” such as “step-ass,” “step-douche,” and “step-dork.” In another such composite construction, “step-führer,” the German loan word “Führer” is also used to describe Chloe’s stepfather as authoritarian. The term “sext/sexting” is an example of blending that can be seen in the dialogue between Max and Juliet about the alleged impropriety of Juliet’s boyfriend Zach. The term sext is a blend of the words “sex” and “text” which is used to describe a sexually explicit photo sent via text message. The term “BFF,” a shortening for “best friends forever,” is used by the character Dana to describe the relationship between Juliet and Victoria. Finally, an example of a shift can be seen in the use of the word “mad” to mean “expert” when the character Alyssa, after dodging an errant football flying toward her head, uses the shift to describe how she is an expert at dodging footballs, claiming that she has “mad skills.”

2.5.5 Texting shorthand in LIS

There are numerous text-message and hand-written exchanges that take place throughout the game. These linguistic exchanges provide authentic examples of how texting shorthand is used in natural conversational contexts. Many of the examples of authentic texting shorthand in LIS take place in simulated mobile text message exchanges between in-game characters. Most of these text message exchanges make use of texting shorthand which retains the phonetic root of the word such as “Mx” for “Max,” “srry” for “sorry,” “thx” for “thanks,” “u” for “you,” and “k” for “okay.” Other text messages exemplify text-only abbreviations for common verbal expressions such as “TTYL (talk to you later)” and “NP (no problem).” Finally, shorthand with no phonetic or orthographic relationship to the abbreviated terminology is shown with the affectionate sign-off expression used at the end of messages with “XOXO (hugs and kisses).”

2.5.6 Cultural referencing in LIS

There are also abundant examples of how cultural referencing functions as nonstandard English language use in LIS. Within the context of social structures, cultural subgroups, also referred to as cliques, are used to categorize people into distinct groups based on any number of broadly defined factors including, but not limited to, interests, behaviors, appearance, socioeconomic status, ethnicity, and abilities (Haenfler, 2013). These cultural subgroups give rise to a plethora of nonstandard language terminology by way of cultural associations with these subgroups as well as in-group language used to signal affiliation or knowledge of the subgroup. As cliques feature prominently within the social fabric of school systems, nonstandard language use with cultural referencing to such cliques is exceedingly common. Set within a US high school, LIS routinely makes use of authentic nonstandard English based on cultural references to cultural subgroups in US culture.

As an example of how linguistic cultural referencing may function in LIS, the narrative structure authentically demonstrates how the terms “jock” and “hipster” can be linguistically applied to broadly define the characteristics of certain characters. In an in-game narrative event, an NPC named Zach accidently throws a football which strikes another NPC named Alyssa in the head. When interacting with Alyssa, she complains to Max about how other students at Blackwell treat her by using the term “jock” as a pejorative. In so doing, the character Alyssa not only is providing linguistic information about the social affiliations of the students but is also distancing herself from the “jock clique” and the host of social characteristics the term stereotypically implies (e.g., poor school performance, social popularity, and athleticism) (Eckert, 1989). If the player tries to interact with Zach, he brushes Max off and uses “hipster” not to describe Max’s fashion sense (Fletcher, 2009) but as a pejorative instead. In this sense, Zach’s use of “hipster” is a linguistically nuanced way of signaling that Max belongs to a different social clique than him and that he therefore wants nothing to do with her.

LIS also makes a number of references to the marijuana subculture in the US through the character Chloe when she smokes marijuana in front of Max. The standard English words “marijuana” and “cannabis” are not used once throughout the entire game. Instead, a host of authentic nonstandard English specific to that subculture is employed. Lying on her bed holding what could be mistaken for a hand-rolled cigarette, Chloe states to Max that she wants to “blaze.” However, the first clue that the substance that Chloe is smoking is not tobacco is the use of the term “blaze,” which can be used in modern English as a verb meaning specifically, “to smoke cannabis” (Merriam-Webster, 2021). Another verb that specifically can mean to smoke cannabis is the verb “toke” which is used when Chloe’s stepfather David questions whether she is “toking up.” A host of nouns referencing the substance cannabis are also employed with “grass,” “weed,” and “joint.” The term “doper,” or frequent drug user, is also employed as a cultural pejorative when David mistakenly thinks that Max brought the cannabis inside his home, referring to her as a “doper.”

2.5.7 Summary of authentic linguistic input

One particularly appealing aspect of the linguistic authenticity that MMORPGs can bring to the foreign language classroom is the high level of exposure to common nonstandard language use, especially considering that most standard language teaching materials do not address the more informal aspects of the TL. With such exposure, language learners will arguably be better prepared for the communicative realities of interacting with both native speakers of the TL as well as the culture of the TL. Single-player RPGs, although lacking the social mechanisms which elicit contact between language learners and native speakers within the context of the game, may nonetheless incorporate authentic linguistic content which features nonstandard language. Single-player RPGs such as LIS therefore may create in-game environments that are highly conducive to providing language learners with exposure to authentic, nonstandard English input.

2.6 Vocabulary acquisition and retention in digital games

Vocabulary strength is one measure of rating a language learner's proficiency in a foreign language. Vocabulary knowledge is required for every interaction that takes place in the TL, and a lack of knowledge in specific TL vocabulary can result in the inability to successfully engage in TL interactions irrespective of proficiency level (Koda, 1989). In other words, TL ability does not always translate into TL comprehension when enough of the vocabulary used in such a linguistic exchange is unknown to the language learner. It has been found that most language learners, irrespective of TL proficiency, do not possess sufficient TL vocabulary knowledge to fully engage in professional and educational tasks in the TL (Laufer, 2000). Moreover, it has been acknowledged that vocabulary plays a vital role in general language development (Barcroft, 2007; Bates, & Goodman, 1997; Cobb & Horst, 2011). As outlined above, exposure to authentic linguistic input is regarded as a key step to developing communicative and cultural competence within the TL. However, simple exposure to authentic linguistic input does not necessarily equate to increases in vocabulary acquisition and retention. The following section will therefore outline how research into the psycholinguistic processes involved with vocabulary acquisition and retention may be applied to SLA contexts in order to achieve desirable pedagogical outcomes; further, the following section will detail why digital games such as MMORPGs and single-player RPGs may be particularly beneficial in eliciting vocabulary acquisition and retention in SLA contexts.

2.6.1 Memory enhancing mechanisms

Much of the research relating to how vocabulary acquisition and retention relate to SLA is based on memory research. With specific regard to how the gameplay mechanics of digital games such as MMORPGs and single-player RPGs create conducive environments for to-be-learned information such as TL vocabulary to be acquired, consolidated, and stored in long-term memory, the following sections will outline six distinct mechanisms: inferential processing, repetition, spaced presentation, consolidation through long-term engagement, multimodal presentation of vocabulary, and vocabulary acquisition through guided play. These memory mechanisms are summarized in Table 2.3. After describing how each mechanism is theorized to affect the acquisition, consolidation, and long-term retention of to-be-learned material, examples will then be provided for how these memory mechanisms are induced through the gameplay mechanics of MMORPGs and single-player RPGs and may function to increase TL vocabulary acquisition.

Table 2.3 Memory mechanisms affecting TL vocabulary acquisition and long-term retention

Memory mechanism	Theoretical effects on TL vocabulary acquisition and long-term retention
Inferential processing	The process of exerting cognitive effort to construct inferences as to the meaning of the unknown TL vocabulary. Sometimes referred to as “effort toward comprehension” (Auble & Franks, 1978) or the “aha effect” (Auble, Franks, & Soraci, 1979), research has identified a robust memory-enhancing effect of utilizing contextual information to form an educated guess, an effect which may also be bolstered through corrective feedback.
Repetition	Widely recognized for its robust effect on memory, repeated exposure to TL vocabulary may result in better long-term retention. Two main hypotheses have been proposed regarding how repetition elicits long-term retention of to-be-learned material: the cumulative-strength hypothesis (i.e., the original memory trace is strengthened through repeated exposure) and the multiple-trace hypothesis (i.e., each exposure creates a unique memory trace which provides multiple cognitive paths to accessing the acquired information) (Hintzman, 2010).
Spaced presentation	Also known as the “spacing effect,” increasing the time intervals between exposures to TL vocabulary may elicit better long-term retention. Hypothesized mechanisms which contribute to better long-term retention of to-be-learned information with spaced presentation (e.g., over multiple study sessions) relative to massed presentation (e.g., in one study session) include: improving semantic encoding through the multiple-trace hypothesis, increasing retrieval/storage strength, and increasing encoding variability (Appleton-Knapp, Bjork, & Wickens, 2005).
Consolidation	The process by which memory traces of TL vocabulary are stabilized in the long-term memory. Long-term engagement with material exposing the learner to the to-be-learned information has been shown to elicit the transformation of acquired memory traces to long-term memory through various processes from repetition to encoding variability (Meeter & Murre, 2004).

Multimodal presentation	The simultaneous presentation of TL vocabulary through various perceptual channels (e.g., auditory, visual, and textual). Multimodal presentation of to-be-learned information is hypothesized to improve acquisition and long-term retention through dual-channel mental representation (Mayer, 2002) and decreased cognitive load on working memory by distributing information along varied perceptual channels (Sweller, 2005).
Guided play	A pedagogical method shown to increase vocabulary acquisition in children. As the process of L2 acquisition (even in adults, albeit at a diminished capacity) is theorized to be largely the same as L1 acquisition, research on the benefits of guided play may be relevant in SLA contexts. Guided play uses semi-structured play sessions whereby children remain largely in control over what and how toys are played with, but adults draw children's attention to learning points such as through scaffolded vocabulary use (Weisberg, Hirsh-Pasek, & Golinkoff, 2013).

2.6.2 Inferential processing and memory

Inferential processing refers to the process of exerting cognitive effort to figure out the answer to or meaning of some unknown piece of information through inferential reasoning and how it affects the retention of the to-be-learned material. Auble and Franks (1978) found that the more effort put into comprehending the meaning of sentences, the better the recall was for the sentence on post-tests. It was therefore concluded that "effort toward comprehension" was positively correlated with enhanced memory recall (Auble & Franks, 1978). The conclusion that effort toward comprehension results in better recall was further elaborated upon in a follow-up study in which it was concluded that effort toward comprehension may be understood as an "aha" experience. In other words, information becomes easier to recall when it undergoes a transformation from a state of incomprehensibility to comprehensibility (Auble, Franks, & Soraci, 1979). Cognitive effort, or effort toward comprehension, is a process of inferential reasoning, whereby comprehended contextual information is employed to draw hypothetical conclusions about unknown information/to-be-learned material. Therefore, the more effort put into the inferential processing of information, the easier it may become to recall that information later. Moreover, the Auble, Franks, and Soraci (1979) follow-up study found that eliciting an "aha moment," whereby the answer to a hypothetical inference was revealed through some form of corrective feedback, enhanced recall of to-be-learned material.

When presented with some type of unknown, to-be-learned material, using contextual information to make an inference about the correct answer, in other words making an educated guess, may result in better recall, especially when the correct answer is provided through a process of corrective feedback. When guessing the answer to some unknown piece of information through the process of inferential reasoning, it goes without saying that incorrect assumptions will be made. The concern that errors made during attempted guesses will become engrained in the long-term memory has been found to be invalid (Kang, Pashler, Cepeda, Rohrer, Carpenter, & Mozer, 2011). In their study, Kang et al. (2011) found that forcing students to guess the answer to an unknown test item did not affect recall performance after corrective feedback was given either immediately or after a delay. Indeed, the very act of guessing incorrectly has been found to have benefits on recalling the incorrect response given for a specific cue in a study that demonstrated a robust effect of inferential processing on memory (Yan, Yu, Garcia, & Bjork, 2014). Furthermore, inferential processing may result in low-confidence responses, even when they are correct. Butler, Karpicke, and Roediger (2008) have found that the process of providing corrective feedback not only improved recall for incorrect responses but also doubled the retention for correct low-confidence responses, emphasizing the potential benefits that inferential processing may have when combined with some form of corrective feedback to improve memory recall.

2.6.3 Repetition and memory

Perhaps one of the most widely recognized and robust effects on memory, simple repetition of to-be-learned material correlates strongly to long-term retention. The exact mechanisms of how the repetition of to-be-learned material correlates to long-term retention is divided into two hypotheses: the cumulative-strength hypothesis and the multiple-trace hypothesis (Hintzman, 2010). The cumulative-strength hypothesis posits that the repetition of to-be-learned material works to strengthen the first memory representation made in the repetition sequence (McClelland, & Chappell, 1998; Murdock, Smith, & Bai, 2001). The multiple-trace hypothesis posits that each repetition of to-be-learned material leaves behind an individual memory trace (Lansdale & Baguley, 2008; Logan, 1988). Gillund and Shiffrin (1984) posit that massed repetitions (i.e., repetition of to-be-learned material within one fixed block of time) create a single memory trace, whereas spaced repetitions (i.e., delayed repetition of to-be-learned material throughout multiple blocks of time) create multiple memory traces. When considering how memory traces interact with the semantic network, multiple memory traces would arguably improve access across the semantic network through creating multiple paths for information recall, which would in turn make spaced repetitions more conducive to long-term retention (Callan & Schweighofer, 2010).

2.6.4 The spacing effect and memory

The spacing effect is indeed a well-researched and robust theory suggesting that spaced presentation elicits better long-term retention relative to massed presentation (Bahrick, Bahrick, Bahrick, & Bahrick, 1993; Dempster, 1988). Outside of the possible connection to better semantic encoding through the creation of multiple memory traces as outlined above, the spaced presentation of information may function to elicit better long-term retention of to-be-learned material by benefiting from the process of retrieval and forgetting. As the interval between presentations of to-be-learned material increases, the retrieval strength of the memory trace decreases. When the retrieval strength of the memory trace becomes too low, the memory can no longer be recalled and is hence forgotten. Ebbinghaus described the process by which the decay function (also known as the "forgetting curve") of retrieval strength is delayed by larger intervals each time the to-be-learned material is restudied, highlighting the importance of spacing out the intervals of restudying (Murdock, 1985). In other words, the more that information is restudied, the more robust retrieval strength becomes, and the longer the intervals between study sessions, the more time the forgetting curve spans.

Bjork (2014) further elaborates upon the concept of retrieval strength through the Theory of Disuse which posits that increases in retrieval strength from the process of restudying to-be-learned material confers more gains to long-term retention for memories with higher starting levels of storage strength (i.e., memories with high levels of inter-associations to other information in the long-term memory). Conversely, memories with higher starting levels of retrieval strength relative to storage strength will elicit gains in retrieval strength after restudying but relatively smaller gains in storage strength. A common experience related to the relationship between retrieval strength and storage strength is "cramming" for examinations; a student can sufficiently remember a large amount of information for a test by a process of studying and restudying the night before, but that information is forgotten just as quickly after the test is over. By focusing on increasing retrieval strength through restudying to-be-learned material in a short amount of time (e.g., massed presentation), retrieval strength may be sufficiently bolstered but at the expense of long-term storage strength, for which relatively smaller gains will be incurred. However, spaced presentations of to-be-learned material are theorized to increase encoding variability by inducing different contextual cues within the learning process, thereby arguably increasing the starting storage strength capacity and long-term retention after retrieval events occur (Appleton-Knapp, Bjork, & Wickens, 2005). The spacing effect may elicit better long-term retention of to-be-learned material by actually encouraging the loss of retrieval strength through longer intervals between restudy, whereby the material must be relearned, which in turn increases encoding variability and thus the durability of long-term memory storage capacity after every retrieval event.

2.6.5 Information consolidation and memory

Similar to the mechanisms that encourage the acquisition of new vocabulary, the consolidation of acquired information into the long-term memory incorporates elements of repetition and encoding variability (Meeter & Murre, 2004). From a pedagogical perspective, encouraging vocabulary acquisition/retention requires that the to-be-learned vocabulary be repeated (ideally through spaced presentation) and accompanied with varied contextual information. It is of no surprise then that extensive reading has been suggested as a way to consolidate acquired TL vocabulary into long-term memory (Pigada & Schmitt, 2006). Extensive reading is a pedagogical strategy that advocates for the long-term engagement with comprehensible TL reading materials (Day, Bamford, Renandya, Jacobs, & Yu, 1998). In so doing, the acquisition of authentic TL vocabulary is elicited through repeated exposure of target vocabulary in the various contexts provided by the source materials. Extensive reading also has the added benefit of an automatic spacing effect due to the syntactic structure of natural language, wherein the varied morphologies of target vocabulary will occur at spaced intervals throughout source materials. Engagement with a pedagogical strategy such as extensive reading in SLA contexts should benefit the long-term retention of TL vocabulary from a psycholinguistic perspective due to spaced repetition and contextual variability to help language learners consolidate acquired TL vocabulary.

2.6.6 Multimodal presentation and memory

Another pedagogical strategy which can be employed to potentially increase the contextual variability within the vocabulary acquisition and consolidation process is the multimodal presentation of linguistic information. Multimodal linguistic information is characterized by the simultaneous presentation of linguistic forms through varied perceptual channels such as the auditory, visual, and textual. Examples of multimodal presentations of linguistic information that may be conducive to SLA processes are multimedia (pictorial/textual) glosses (Yanguas, 2009), graphic novels (Wrobetz, 2018), subtitled media (Borrás & Lafayette, 1994; Frumuselu, De Maeyer, Donche, & Plana, 2015), and digital games (Newon, 2011; Zheng, Bischoff, & Gilliland, 2015). The use of multimodal linguistic information to improve retention of TL vocabulary is supported by Mayer's Multimedia Theory which posits that mental representations built out of both textual and visual information (i.e., the dual-channel assumption) are more conducive to the learning process (Mayer, 2002). Furthermore, Sweller's Cognitive Load Theory suggests that the multimodal presentation of linguistic information is more conducive to long-term memory consolidation through a reduction of cognitive load on the working memory elicited by the distribution of information along varied perceptual channels (Sweller, 2005). The manner in which varied modalities of information work together may be particularly effective in reducing cognitive load in SLA contexts, whereby the visual and textual representations of linguistic forms work to support what might otherwise be incomprehensible auditory input in the TL (Mayer, Lee, & Peebles, 2014).

2.6.7 Guided play and memory

Ludic interactions with language in a semi-structured, socially interactive environment called "guided play" have been shown to increase vocabulary acquisition and retention in early childhood (Weisberg, Hirsh-Pasek, & Golinkoff, 2013). Long (1990) suggests that the processes involved in L1 acquisition in early childhood may remain largely the same later in life for L2 acquisition, albeit to a diminished degree. Should the maturational constraints seen in SLA in which the faculties for language acquisition, as Long suggests, become atrophied due to disuse, then analyzing how L1 vocabulary acquisition functions in early childhood may have broad pedagogical applications in SLA contexts. Guided play is a pedagogical approach in early education which involves allowing children to direct how a playful activity will evolve but while providing adult interaction to reinforce learning points such as scaffolded target vocabulary use (Weisberg, Hirsh-Pasek, & Golinkoff, 2013). Guided play has been shown to achieve better pedagogical outcomes for vocabulary acquisition and retention than direct instruction (Stipek, Feiler, Daniels, & Milburn, 1995) or child-directed play free from scaffolded adult interaction (i.e., free play) (Chien et al., 2010). Learning through playful interactions such as guided play may have pedagogical implications for SLA by helping to improve well-established, quasi-playful pedagogical practices such as conversational role-playing (Félix-Brasdefer, 2018). In particular, by using the principles of guided play and allowing SLA conversational role-playing to be more student directed and teacher guided (in contrast to reading a script from a book), better pedagogical outcomes for vocabulary acquisition and retention may be achieved, especially when ludic elements are integrated into the activity.

2.7 Vocabulary acquisition and retention in MMORPGs

A large volume of research in the relevant literature suggests that MMORPGs played in SLA contexts are beneficial to TL vocabulary acquisition and retention (Li, 2019b). In order to analyze how some of the underlying gameplay mechanics in MMORPGs work to elicit TL vocabulary acquisition and retention, the six memory-enhancing mechanisms introduced above in Table 2.3 will be applied to specific examples of MMORPG gameplay to add further evidence to the conceptual framework of the pedagogical efficacy of MMORPGs in SLA contexts.

2.7.1 Inferential processing in MMORPGs

Identifying how MMORPG gameplay mechanics allow language learners to use the context of the game to make inferences about the meaning of in-game vocabulary may aid in understanding how conducive MMORPGs are to TL vocabulary acquisition/retention. For example, if a language learner hypothetically comes across an unknown TL word during gameplay such as "citadel," a number of inferences may be made as to what the correct pronunciation and meaning in English might be. Prior knowledge of English pronunciation and spelling will certainly be drawn upon, perhaps to infer that the "c" should either be pronounced as an "s" as in "century" or as a "k" as in "castle." Other contextual cues might be used to infer the meaning such as the other words in a hypothetical sentence, "You must traverse to the citadel in the west" perhaps to arrive at the conclusion that a citadel is a location such as a city or a structure. Should the language learner have to communicate this in-game information to other players in order to progress in the game, then recall of both the pronunciation and meaning to the word citadel will theoretically be enhanced, especially when corrective feedback is provided within the framework of the social gameplay mechanics.

2.7.2 Repetition in MMORPGs

Another memory-enhancing mechanism that may function to increase TL vocabulary acquisition/retention within the gameplay mechanics of MMORPGs is repetition (Steinkuehler, 2004). Moreover, the repetition of in-game vocabulary will necessarily occur embedded with varied contextual cues as MMORPG gameplay mechanics will create new contexts as the player progresses through the game. Not only have language learners who use MMORPGs as language learning resources identified the repetition and noticing of frequency for in-game vocabulary as a strategy for improving TL vocabulary strength (Bytheway, 2015), but the repetition of in-game vocabulary is also thought to lead to a state of automaticity through "overlearning" (Coxhead & Bytheway, 2015). That repetition can lead to a state of automaticity over time is well supported by the concept of "repetition priming" which is theorized to be a cumulative process involving the improvement of implicit memory (i.e., automatized memory) over explicit memory (i.e., intentionally recalled memory) (Rugg, Mark, Walla, Schloerscheidt, Birch, & Allan, 1998). From an SLA perspective, repetition priming may also improve the automaticity of morphologically related words through the cognitive organization of lexical items into groups of word components such as stem morphemes (Fowler, Napps, & Feldman, 1985).

An example of how this repetition process may function within the context of an MMORPG may be found in an in-game status effect common to RPGs known as "petrification" whereby an enemy, ally, or one's own character turns to stone. As the context of the word "petrification" varies within the game, so too will its morphology, ranging from verb conjugations (petrify, petrifies, petrified) to the noun (petrification). As a language learner is continually exposed to the word "petrification" within the context of gameplay, the long-term retention of not only the word itself but also the automaticity of the stem morpheme "petr(a)-" may be enhanced. This state may also theoretically impact the cognitive lexical organization of words that may not have appeared during gameplay but are nonetheless morphologically related such as "petroleum."

2.7.3 Spaced gaming in MMORPGs

The acquisition and retention of TL vocabulary from MMORPGs may also benefit from the spaced repetition of in-game vocabulary. This in-game spacing between the presentation of vocabulary may come as the result of intervals between gaming sessions, which has been shown in a study analyzing WoW players' gaming hours to alternate between on and off periods shorter than five days for approximately 80% of players (Tarng, Chen, & Huang, 2008). However, the spaced presentation of in-game vocabulary may also be thought of as being an independent function of the gameplay mechanics unique to both MMORPGs and single-player RPGs, namely through a process of linguistic up-regulation and down-regulation (Wrobetz, in press). Although many digital RPGs make heavy use of linguistic elements within their gameplay mechanics (e.g., through storylines), there is also heavy use of gameplay mechanics which do not require the use of any language whatsoever (e.g., fighting enemies). Many digital RPGs are first and foremost commercial products following a long line of prior developments that have successfully maintained market relevance. Controlling the strategic development of a character to successfully battle through progressively more and more difficult in-game challenges is a common core element to gameplay. The linguistic elements of gameplay, story included, may arguably be thought of as mechanisms which serve as a type of packaging to deliver the core gameplay mechanics to the customer, namely leveling one's character up to engage with the battle system.

In MMORPGs, the intervals of gameplay during which the player is less attuned to the linguistic elements of gameplay, such as when engaged with the battle system, may be described as intervals of linguistic down-regulation. Conversely, linguistic up-regulation would describe intervals of gameplay during which the player would be more attuned to linguistic information, such as when engaged with story elements or in conversation with other players or NPCs. As the player cycles through these intervals of linguistic down-regulation and up-regulation, an in-built mechanism for spacing presentations of TL vocabulary emerges. From a psycholinguistic perspective, the presence of such a linguistic spacing mechanism would encourage long-term retention of any TL vocabulary acquired through gameplay and would arguably add another theoretical explanation as to how engagement with MMORPGs in SLA contexts are eliciting gains in TL vocabulary acquisition and long-term retention.

2.7.4 Information consolidation in MMORPGs

The consolidation of TL vocabulary through long-term engagement with MMORPGs is yet another mechanism that supports their use in SLA contexts. In particular, MMORPGs may also work to consolidate TL vocabulary acquired during gameplay but also TL vocabulary acquired outside gameplay, namely from educational contexts. In a study analyzing the effects of extramural digital gameplay on EFL performance in Swedish elementary schools, Sylvén and Sundqvist (2012) found a positive correlation between higher frequencies of digital gameplay outside the classroom, especially with MMORPGs, and higher scores on standardized EFL proficiency tests measuring vocabulary knowledge, reading comprehension, and listening comprehension. These findings may provide evidence to support the argument that the process of TL vocabulary consolidation that occurs as a result of playing MMORPGs in SLA contexts is not limited to in-game vocabulary and may extend to TL vocabulary encountered outside the context of the game. This TL vocabulary consolidation through gameplay may be taking place as a result of in-game repetition of TL words which are taught in class or as a result of improved morphological automaticity from the in-game repetition priming of TL vocabulary on standardized tests. The indication that higher levels of engagement with MMORPGs translates to concrete gains in L2 proficiency in contexts outside of gameplay supports the conjecture that the gameplay mechanics of MMORPGs elicit long-term retention through extensive exposure to relevant TL vocabulary.

2.7.5 Multimodal presentation in MMORPGs

The multimodal presentation of linguistic information is standard for most digital games, MMORPGs being no exception. The multimodal presentation of linguistic information in MMORPGs includes the simultaneous presentation of voice-acted content, subtitles for voice-acted in-game content, in-game text and dialogue, voice chat, text chat, and visual renderings of characters and environments. However, the key feature of multimodality that sets digital games such as MMORPGs and single-player RPGs apart from other forms of multimodal media such as subtitled film is the ability to interact with the in-game environment. In contrast to film where the language learner may be described as a passive observer to the multimodal presentation of linguistic information, digital RPGs put the language learner in control of the manner and frequency of linguistic exposure. The additional layer of multimodality in MMORPGs provided by the interactivity of the medium of digital gaming would arguably enhance the theoretical benefits to TL vocabulary acquisition and long-term retention elicited by the simultaneous presentation of in-game linguistic information in visual, textual, and auditory modalities as may be supported by Multimedia Theory and Cognitive Load Theory (Mayer, 2005; Sweller, 2011).

2.7.6 Guided play in MMORPGs

The ludic elements combined with the objective-based gameplay mechanics in MMORPGs may also constitute another mechanism to elicit higher rates of TL vocabulary acquisition and retention when used in SLA contexts. When compared to the pedagogical methodology of guided play, the similarities to the gameplay mechanics of MMORPGs draw into question whether the aspect of semi-freeform play may be contributing to SLA gains from MMORPG gameplay and whether there is an age limit to the benefits that guided play brings to bear on the vocabulary acquisition process. Although much of the research on how play and learning, including language development, are interrelated is focused on early childhood development (Roskos, Christie, Widman, & Holding, 2010), when considering that the learning process of SLA may occur at any point in one's life, such research may shed light on how playful activities may be beneficial to language acquisition regardless of age. Paul Gee (2003a) even goes so far as to argue that human learning in and of itself is a playful process and that when play stops, so too does learning.

The definition of play being used in the following analysis is based on relevant research (Burghardt, 2011), namely that play is different from other activities because it has no specific application to survival, involves pretend actions that are exaggerated from real actions, and is done voluntarily by the participant. Regarding how play and language development are interrelated, Weisberg et al. (Weisberg, Zosh, Hirsh-Pasek, & Golinkoff, 2013) identify four main areas that play and language development are interrelated in Table

2.4. Guided play (i.e., child-directed play with semi-structured adult intervention) has been shown in research to elicit higher rates of vocabulary acquisition relative to free play and directed play (Weisberg, Zosh, Hirsh-Pasek, & Golinkoff, 2013), potentially due to the added input provided by a linguistically more capable participant (i.e., the adult). When comparing the scenario of a child interacting with a linguistically more capable peer in a symbolic social environment over which they can exert a large amount of control to that of a language learner interacting with linguistically more capable peers (e.g., native speakers or highly proficient L2 speakers) in a symbolic social environment (e.g., the virtual world of the MMORPG) over which they can exert a large amount of control (e.g., character development path), voluntary participation in MMORPGs for fun may arguably constitute a type of guided play in SLA contexts.

Table 2.4 Relationships between play and language development

Identified interrelationship	Proposed mechanism
Play and symbolism	Play involves symbolic substitutions for real objects much in the same way that words are symbolic representations of real objects.
Play and social interaction	Play involves social interactions and provides contextualized tasks for the utilization of language.
Play and linguistic input	There is a large amount of linguistic input involved in playful activities.
Play and personalization	Since the participants in the playful activity are in control of the activity, they are engaged with the activity and follow what interests them.

Note. Summarized from a study by Weisberg, Zosh, Hirsh-Pasek, and Golinkoff (2013)

2.7.7 Summary of memory mechanisms in MMORPGs

As has been outlined above, the gameplay mechanics of MMORPGs create an environment conducive to vocabulary acquisition and retention in SLA contexts through the use of six mechanisms which are associated with enhanced recall of acquired vocabulary (see Table 2.3). MMORPGs require players to process unknown target vocabulary inferentially by using the context of the in-game objectives. The correct meaning of inferences made for unknown target vocabulary are then negotiated with other players or subject to a process of corrective feedback. Acquired TL vocabulary are also repeated in various contexts throughout the game which may prime the learner to learn morphologically similar TL vocabulary. Intervals of linguistic up-regulation and down-regulation create in-game spacings between the presentation of TL vocabulary which may increase long-term retention and encoding variability. High frequency engagement with MMORPGs has been shown to be correlated to the consolidation of TL vocabulary in out-of-game contexts. The multimodal presentation of linguistic information includes voice-acted content, subtitles, in-game text and dialogue, voice chat, text chat, and visualizations. Finally, the player-directed, voluntary progression through a structured digital gameplay environment that enables social interactions with linguistically more capable peers resembles the pedagogical method of guided play, which has been shown to increase vocabulary acquisition and language development. Many of the mechanisms acting on enhancing vocabulary recall in MMORPGs benefit from the social element of the gameplay mechanics. The following section will delve into how these recall-enhancing mechanisms may or may not function in single-player RPGs without social gameplay mechanics.

2.8 Vocabulary acquisition and retention in single-player RPGs

It bears repeating that the key fundamental difference in gameplay mechanics between MMORPGs and single-player RPGs are the social gameplay mechanics, yet the underlying principle of RPG gameplay remains largely the same. Whereas some of the theoretical benefits to TL vocabulary acquisition and retention in MMORPGs hinge on the social elements of gameplay, many of the non-social gameplay mechanics that overlap with MMORPGs may also benefit the vocabulary acquisition process in SLA contexts. The following sections will analyze how the six memory mechanisms (see Table 2.3) function in single-player RPGs.

2.8.1 Inferential processing in single-player RPGs

Returning to the example of how inferential processing may help to bolster TL vocabulary acquisition/retention when combined with corrective feedback processes, single-player RPGs may benefit from similar psycholinguistic mechanisms as observed in MMORPGs. Given the same scenario described in the analysis of inferential processing in MMORPGs in which the word "citadel" in the sentence, "You must traverse to the citadel in the west" is not understood by a language learner playing a single-player RPG, the same inferences may be made by a language learner player if the information needs to be understood in order to progress in the game (i.e., the player must travel to the citadel to complete an in-game objective). However, the effectiveness of this inferential processing will be entirely dependent on the gameplay mechanics of the single-player RPG to elicit corrective feedback and allow for negotiation of meaning to occur within the in-game environment. However, in contrast to MMORPGs, single-player RPGs must elicit such corrections/negotiations without leveraging social interactions. These non-social corrective gameplay mechanics which help move the player through in-game objectives are common in many single-player RPGs.

2.8.2 Repetition in single-player RPGs

One of the mechanisms in single-player RPGs for TL vocabulary acquisition/retention that functions identically to those in MMORPGs is the simple process of repetition. Returning to the example from MMORPGs of how the repetition of the common status effect "petrification" in varied circumstances and morphologies throughout gameplay may benefit TL vocabulary acquisition/retention, the same process of repetition priming may be observed in single-player RPGs. In the single-player RPG *Final Fantasy IX* (Square, 2000), not only does the common status effect of petrification come up during battle sequences but the status effect is also discussed periodically throughout the game as a part of the narrative. At the beginning of the game in a location known as "Evil Forest," a quasi-playable character named Blank becomes petrified along with the forest. Within the context of the game, an item known as "soft" cures petrification. However, due to the scale and severity of the forest's petrification, the player embarks on a prolonged quest to locate a more potent version of the potion known as "supersoft." Dialogue regarding the quest for the supersoft comes up periodically throughout the game and provides varied contexts and morphological variations for the word. Just as with MMORPGs, the repetition priming of in-game vocabulary like "petrification" in single-player RPGs may benefit not only the acquisition and retention of the stem morphemes but also the comprehension of morphologically related vocabulary.

2.8.3 Spaced presentation in single-player RPGs

Repetition priming in single-player RPGs also benefits from similar gameplay mechanics observed in MMOPRGs which allow for the spaced presentation of in-game vocabulary. One manner in which the spacing of TL vocabulary presentation is accomplished is through the linguistic up- and down-regulation throughout the game. One example of linguistic down-regulation common to both MMORPGs and single-player RPGs occurs during periods of "grinding" or otherwise engagement with in-game activities that do not progress the story, but which contribute to the enhancement of a player's character (e.g., acquiring new in-game abilities) or focus on the acquisition of optional in-game items such as equipment or magic spells (Zagal & Altizer, 2014). Grinding usually necessitates the repetition of specific in-game tasks which tend to be linguistically down-regulated, such as fighting enemies over and over again so as to increase the level of one's character, acquire a new in-game ability, or earn in-game money to purchase an in-game item. In a study examining how TL vocabulary acquisition and retention functions with the single-player RPG *Final Fantasy IX*, it was found that intervals of down-regulation that came about as a result of grinding significantly decreased the rate at which in-game vocabulary acquisition occurred relative to intervals of in-game linguistic up-regulation (Wrobetz, in press). From a psycholinguistic perspective, gameplay mechanics which oscillate between intervals of linguistic up- and down-regulation may be conducive for enhancing the recall of acquired TL vocabulary by way of spacing the presentation of linguistic information throughout the total gameplay interval.

2.8.4 Information consolidation in single-player RPGs

The extensive reading opportunities and language resources offered by continuous long-term engagement with story-centered single-player RPGs may also provide long-term consolidation of acquired language in contexts outside of gameplay. In Sylvén and Sundqvist's study (2012) on the educational impact of extramural digital gaming habits of Swedish elementary school students, students who played MMORPGs in English outperformed both students who did not engage in extramural digital gameplay as well as students who played offline single-player games such as *The Sims* (Maxis, 2000) on national standardized testing. Based on this data, the argument could therefore be made that MMORPG gameplay mechanics result in superior long-term consolidation of language acquired through gameplay relative to offline single-player games. However, such an argument would belie the fact that engagement with any genre of digital game, even offline single-player games, conferred SLA benefits in the Sylvén and Sundqvist study and that the widely varying gameplay mechanics of offline single-player games may achieve similarly disparate results for SLA in the absence of the social gameplay mechanics observed in online digital games. Returning to the argument regarding the pedagogical applicability of commercial digital games, not all games in the umbrella genre of "offline single-player games" will achieve equal results, and it is therefore imperative to identify how the gameplay mechanics of single-player RPGs differ from other offline single-player games for SLA purposes.

One key feature that may provide evidence that single-player RPGs are more beneficial in SLA contexts relative to other genres of offline single-player games would be the amount of linguistic up-regulation built into the gameplay mechanics. In other words, single-player RPGs, especially those whose gameplay revolves around the unfolding of a central narrative, not only make use of high frequencies of linguistic input but also require that that linguistic input is comprehended in order to progress through the game. The offline single-player game *The Sims* listed in the Sylvén and Sundqvist study (2012) is a simulation game which is relatively linguistically down-regulated when compared to more story-centered single-player RPGs such as the Final Fantasy series. Not only is there no central plot per se within *The Sims* which must be comprehended in order to progress through the game, but the characters in the game do not even speak an intelligible language, using instead an in-game pseudo language known as "Simlish" (Portnow, 2011). With the absence of a central plot which generates in-game objectives and greatly reduced exposure to authentic language within the context of the game by essentially relegating all authentic language use to menu controls, simulation games such as *The Sims* would, from both sociocultural and psycholinguistic perspectives, appear less applicable to SLA contexts than offline single-player games with higher degrees of in-game linguistic up-regulation.

In a study examining two non-native English speaker players' communal interactions with the story-centered single-player RPG *Final Fantasy X* (Square Product Development Division 1, 2001), it was found that English vocabulary acquired through gameplay transferred to out-of-game contexts (Piirainen–Marsh & Tainio, 2009). Although not directly comparable to the Sylvén and Sunqvist (2012) results from either the MMORPG or the offline single-player gaming group, story-centered single-player RPGs such as *Final Fantasy*

X that make use of intervals of in-game linguistic up-regulation offer a more conducive environment for SLA processes from a psycholinguistic perspective. *Final Fantasy X* exposes language learner players to extensive TL language resources through an average of 75 hours of total gameplay (HowLongToBeat, 2021) which oscillates between intervals of linguistic up- and down-regulation. During the intervals of linguistic up-regulation, the central plot generates in-game objectives which require the player to comprehend the linguistic input. Therefore, gameplay mechanics which not only provide long-term exposure to authentic TL input but also necessitate the comprehension of that input may be an effective predictor of the ability for how well digital games elicit long-term consolidation of acquired vocabulary than the simple presence or absence of social gameplay mechanics.

2.8.5 Multimodal presentation in single-player RPGs

Another factor that makes a single-player RPG such as *Final Fantasy X* a more applicable game to SLA contexts relative to an offline single-player simulation game such as *The Sims* is the full integration of voice-acted dialogue. Along with visualizations, text-based dialogue, and subtitles, the extra layer of linguistic input afforded by voice-acted dialogue arguably creates a more conducive environment for TL vocabulary acquisition/retention. In contrast to games that use in-game pseudo languages such as Simlish in *The Sims*, single-player digital games which offer a fully voice-acted in-game experience may have improved functionality in SLA contexts by reducing cognitive load across multiple perceptual channels for linguistic processing (Joseph & Uther, 2009; Mayer, 2002; Mayer, 2005; Mayer, Lee, & Peebles, 2014). As full integration of voice-acted dialogue has become more and more commonplace amongst high-budget single-player RPG titles such as *Final Fantasy X* and LIS, so too arises the argument that single-player titles featuring such gameplay mechanics have the ability to elicit enhanced TL vocabulary acquisition/retention based on the multimodal presentation of linguistic input. However, any theorized benefit to TL vocabulary acquisition/retention elicited by such multimodal presentation of linguistic input is entirely dependent on the quantity and quality of the multimodal in-game content.

2.8.6 Guided play in single-player RPGs

The ludic elements of single-player RPGs in conjunction with the player-directed interactions within the semi-structured in-game environment may elicit enhanced vocabulary acquisition and retention relative to non-interactive teaching material over which the learner exerts no control. The pedagogical method of guided play has been theorized to contribute to enhanced early-childhood vocabulary acquisition relative to free and directed play by allowing children to create an engaging ludic scenario with scaffolded feedback from linguistically more capable peers (i.e., the adult). The gameplay mechanics of single-player RPGs mimic the guided play principles by allowing the player to direct both the development of playable in-game characters and their interactions with the in-game environment which scaffolds in difficulty as the player progresses through structured gameplay. As research within the field of SLA suggests that the same mechanisms are involved in both L1 and L2 acquisition, looking at pedagogical methods which enhance L1 vocabulary acquisition in children can inform pedagogical practice in SLA contexts as well (Unsworth, 2007). Single-player RPG gameplay mechanics which require the comprehension of in-game linguistic input in order to perform increasingly difficult in-game tasks may therefore elicit enhanced vocabulary acquisition and retention of TL vocabulary in a similar manner as has been observed with the guided play methodology irrespective of age.

2.8.7 Summary of memory mechanisms in single-player RPGs

As has been outlined above, the gameplay mechanics of single-player RPGs create an environment conducive to vocabulary acquisition and retention in SLA contexts through the use of six mechanisms (see Table 2.3) which are associated with enhanced recall of acquired vocabulary. Single-player RPGs require players to process unknown target vocabulary inferentially by using the context of the in-game objectives. The correct meaning of inferences made for unknown target vocabulary are then negotiated with the in-game environment or subject to a process of corrective feedback provided by the gameplay mechanics. Acquired TL vocabulary are also repeated in various contexts throughout the game which may prime the learner to learn morphologically similar TL vocabulary. Intervals of linguistic up-regulation and down-regulation create in-game spacings between the presentation of TL vocabulary which may increase long-term retention and encoding variability. Long-term engagement with single-player RPGs which require on some fundamental level the comprehension of in-game linguistic input has been shown to be correlated to the consolidation of TL vocabulary in out-of-game contexts (Piirainen–Marsh & Tainio, 2009; Wrobetz, in press). The multimodal presentation of linguistic information includes voice-acted content, subtitles, in-game text and dialogue, and visualizations. Finally, the player-directed, voluntary progression through a structured digital gameplay environment that encourages active engagement with in-game linguistic input mimics the pedagogical method of guided play, which has been shown to increase vocabulary acquisition and language development. The following sections will delve into how these recall-enhancing mechanisms may or may not function in LIS.

2.9 Vocabulary acquisition and retention in LIS

With specific regard to how the gameplay mechanics of LIS create conducive environments for TL vocabulary to be acquired, consolidated, and stored in long-term memory, the following sections will use the six memory mechanisms (see Table 2.3) introduced above to frame the argument that LIS may be pedagogically effective in SLA contexts. The following sections will introduce specific examples from LIS for each of the six memory mechanisms to illustrate how the gameplay mechanics specific to LIS may function to increase TL vocabulary acquisition and long-term retention.

2.9.1 Inferential processing in LIS

Much like many other single-player RPGs, LIS may provide a gaming experience that is conducive to vocabulary acquisition and retention in SLA contexts through non-social gameplay mechanics that encourage inferential processing of unknown TL vocabulary necessary to complete in-game objectives. Furthermore, both the corrective mechanics and interactive in-game environment may improve recall for TL vocabulary. Returning to the LIS in-game objective in which the main character Max must retrieve her friend Warren's flash drive from her dormitory room, the player is confronted with a number of sub-tasks that stand in the way of completing the larger in-game objective of fetching the flash drive from her room. One of these in-game tasks involves forcing Max's rival Victoria and her two friends to move away from the entry of the dormitory so that Max can enter. While trying to figure out how to solve this in-game sub-task, a language learner playing LIS in English will arguably have to infer the meaning of a number of linguistic clues, negotiate the meaning of unknown words by exploring the in-game environment, and receive explicit corrective feedback from the game should the player have incorrectly understood what must be done in order to complete this sub-task.

As the entrance to the dormitories is blocked by Victoria and two of her friends, the player is forced to enter into a linguistic exchange with the group of girls. After a tense exchange between Max and Victoria, it is revealed that Victoria and her friends have no intention of moving out of the way to let Max into the dormitory. Max proceeds to establish the sub-task parameters with an inner monologue stating that she will find a way to get Victoria and her friends to move. There are a number of potentially difficult English expressions for language learners that come up during the exchange which may prompt the player to make inferences about the meaning of these exchanges. Moreover, as the player cannot progress any further in the game until the sub-task is complete, a certain amount of effort must be made by the player to understand these linguistic exchanges and dialogue prompts. The potential for effort toward comprehension in the scenario described above will ultimately result in unknown TL vocabulary that have undergone inferential processing to be either confirmed or corrected by the gameplay mechanics.

The corrective mechanics may elicit explicit corrective feedback, which will force the player to use the time reversal function or encourage the player to negotiate what must be done through the exploration of the interactive environment. Should a language learner have miscomprehended the in-game prompts to force Victoria and her friends to move, then the player may wander around the area looking for another way into the dormitory. The gameplay mechanics do, however, provide visual clues as to what must be done in order to move Victoria and her friends out of the way. The groundskeeper, Samuel, slowly walks towards a ladder perched against some scaffolding set up around the stoop upon which Victoria and her friends are sitting. Upon reaching the ladder, Samuel picks up a bucket of paint, climbs up the ladder, and hangs the bucket of paint off a hook. No matter where the character is exploring, as soon as Samuel hangs the bucket off the hook, the screen will show that animated sequence to the player. This is the first game-prompted clue that the player must do something with the paint bucket in order to progress the game forward. If the player still does not understand what must be done, the gameplay mechanics step in with explicit corrective feedback, moving to a black and white still-frame of the hanging paint bucket, stating that something has been missed. This screen forces the player to rewind to the point just after the dialogue between Max and Victoria ended. Once the player’s attention has been sufficiently drawn to the paint bucket, the player may begin to negotiate the meaning of what has to be done by interacting with the environment and unveiling player-prompted clues.

Upon being forced to rewind by the corrective gameplay mechanics, the player may choose to examine the paint bucket before Samuel reaches it. Two options appear upon examining the paint bucket: “look” and “tamper.” Choosing to look at the bucket reveals a player-prompted clue as to what must be done to solve this specific sub-task as Max explicitly states that the paint bucket’s proximity to Victoria is a part of the solution. Should the word “tamper” be unknown to a language learner player, specific contextual information provided by the in-game scenario may work to encourage inferential processing. Defined as the act of “exerting secret or corrupt influence upon someone or something” (Merriam-Webster, 2021), tampering with the paint bucket so as to make it fall on Victoria and her friends requires that Max carry out the action in secret. Indeed, as the groundskeeper Samuel is slowly walking towards the ladder and paint bucket, if the player tries to tamper with the handle of the paint bucket while Samuel is nearby, Max will state that the groundskeeper is too close to execute the “tamper” command. The player will then have to use the time reversal mechanics to ensure that Samuel is far enough away to allow Max to covertly tamper with the handle of the paint bucket. Once all the necessary conditions have been met and Max tampers with the handle of the paint bucket so as to make it fall when Samuel hangs it off the hook, a statement from Max informs the player that the actions taken leading up to the tampering have been correct.

All of the player’s interactions with the in-game environment which have led to inferences about what must be done are either confirmed or corrected by what happens when Samuel hangs the tampered-with paint bucket. If the player did not explore the environment to gather player-prompted linguistic clues as to how to solve the sub-task beyond examining the paint bucket, then the paint bucket falls to the ground but falls short of affecting Victoria and her friends. In such a scenario, the task will remain unsolved, and the player will be unable to progress. From the standpoint of a language learner player who may not fully comprehend the language used in the linguistic clues, miscomprehension of the linguistic input may be the cause for failing to progress through the game. For example, by examining the sprinkler, the linguistic clue of "cranking the sprinkler up" so Victoria and her friends will "beat it" is provided to inform the player that adjusting the pressure in the sprinkler is a necessary step to solve the sub-task. The inclusion of the nonstandard terminology "crank up" and "beat it" arguably creates a linguistic barrier to comprehending the nature of the clue, namely that the player must increase the sprinkler pressure in order to get Victoria and her friends to move to a position where the falling paint will splash on them. Should the player still not be able to complete the in-game task, corrective gameplay mechanics will guide the player to the correct solution. Following the corrective gameplay mechanics will eventually guide the player to the correct sequence of events, whereupon the paint falls on Victoria. The player may then enter into the dormitory to continue with the original in-game objective of retrieving Warren's flash drive.

In a single-player RPG such as LIS which has high levels of linguistic up-regulation requiring the player to comprehend a number of linguistic clues within the in-game dialogue, a number of factors may force a language learner player to make inferences about what must be done to solve in-game tasks. In the example above, the high frequency of nonstandard English terminology coupled with the multi-layered nature of the in-game task itself (e.g., using the sprinkler and tampering with the paint bucket) will arguably impose barriers to full comprehension as to what must be done in order to progress further in the game; in other words, the nature of the in-game tasks in LIS will arguably force language learner players to put effort into comprehension and create an environment conducive to inferential processing. This effort toward comprehension that the language learner will theoretically have to go through will ultimately be confirmed by way of successful completion of the in-game task through a process of correction/negotiation. These corrective mechanics are accomplished in

LIS by game-prompted explicit feedback (e.g., forced reverse), game-prompted clues (e.g., showing Samuel and his paint bucket), and player-prompted clues stemming from the exploration of the interactive in-game environment (e.g., investigating the paint bucket or sprinkler). From a psycholinguistic perspective, such gameplay mechanics will enhance the acquisition/retention of any TL vocabulary that has undergone inferential processing and corrective feedback/negotiation.

2.9.2 Repetition in LIS

Just as with other single-player RPGs, LIS may also enhance recall of acquired TL vocabulary through the repetition of in-game vocabulary. In LIS, vocabulary repetition occurs in nearly identical manners as with other digital RPGs, namely as integral components of in-game dialogue or as frequently used in-game controls. For example, the word "flash drive" comes up on numerous occasions as an integral component of the in-game objective. Similarly, the menu item "rewind" is selected numerous times throughout the game as a part of the time reversal mechanics to enable the player to redo in-game scenarios and progress through the game. However, there are also gameplay mechanics which are specific to LIS which may work to increase the repetition of in-game vocabulary outside of the above-stated scenarios. The time reversal mechanics in LIS, the SLA functions of which are summarized in Table **2.5**, work to create an in-game environment conducive to repetition by exposing the player to the same in-game dialogue sequences over and over again. One of the functions of the time reversal mechanics, as has been described in earlier sections, is to allow the player to use information gathered from in-game social exchanges/in-game tasks and change their outcome. There are also two other functions of the time reversal mechanics which are frequently used throughout the game and work to create an environment conducive to the repetition of in-game vocabulary.

Table 2.5 Uses of time reversal mechanics in LIS

In-game use	In-game function	Hypothesized effect on TL vocabulary acquisition/retention
Script choices/In-game tasks	Allows the player to redo social exchanges with NPCs and change their selection to other script choices available. Similarly allows players to redo in-game actions performed during in-game tasks. Occasionally, using the time reversal mechanics after being exposed to relevant information in social exchanges will create new script choices upon reversal.	The player's script choices may result in different dialogue sequences and may act as a quasi-productive linguistic tool for the player to role-play in the TL. When used to solve in-game tasks, the time reversal function will provide the player with repeated exposure to TL vocabulary relevant to solving the in-game task. Some social exchanges require the input of specific script choices. Incorrect script choices will impede progress through the game, therefore the time reversal mechanics act as a mechanism for corrective feedback and elicits repeated exposure to TL vocabulary relevant to solving such script choice puzzles.
Action sequences	Allows the player to redo in-game actions during a timed action sequence. If specific in-game actions are not done within the specified amount of time, the player cannot progress through the game.	While performing the necessary in-game actions during timed action sequences, in-game dialogue may still be occurring between other NPCs. If the player requires several time reversals in order to successfully complete the action sequence, the player will be repeatedly exposed to the same TL vocabulary relevant to the context of the action sequence.
Narrative choices	Allows the player to redo narrative choices within the game. In contrast to normal script choices, narrative choices will impact how the narrative structure of the game plays out.	As narrative choices impact how the narrative structure unfolds, the player may be incentivized to redo the narrative choices to unlock different narrative paths, thus increasing the exposure to relevant TL vocabulary.

2.9.3 Action sequences in LIS

As opposed to other digital RPGs, the choose-your-own-adventure genre of LIS relies much less on the development of playable characters by way of accumulating in-game money and experience points through a battle system. Indeed, as character development in LIS revolves around how in-game choices affect the unfolding of the narrative, there is little need for the grinding associated with other digital RPGs (e.g., continuously battling enemies to collect money or experience points). The battle systems common in most digital RPGs, while certainly being a source of in-game repetition, are significantly linguistically down-regulated as the player's focus is shifted from the comprehension of in-game dialogue to pushing the correct buttons to defeat enemies. While LIS does not make use of a battle system, there are action sequences in which the player must perform a specific in-game action within a short interval of time in order to progress in the game. Just as with the battle systems in other digital RPGs, the action sequences in LIS work to shift the player's attention away from the comprehension of in-game dialogue to pushing the correct buttons to quickly carry out the in-game task. However, as these action sequences require the player to carry out a specific in-game task within a short interval of time, the failure to accomplish the in-game task will prompt the player to use the time reversal mechanics as many times as is required until the task is accomplished correctly. Therefore, although these action sequences arguably represent the highest degree of linguistic down-regulation in LIS, the time reversal mechanics may work to increase the repetition of in-game vocabulary while carrying out these action sequences.

One example of such an action sequence in LIS takes place when Max must save Chloe from being shot by the character Nathan in the girls' restroom. In order to save Chloe from getting shot by Nathan, the player is prompted to pull the fire alarm in the girls' bathroom. There are numerous steps involved in pulling the fire alarm and if the player does not execute the actions quickly enough, Chloe is shot and the player is forced to reverse time in order to attempt the in-game task again. The total number of required rewinds depends entirely on how fast the player is able to perform the required steps during this action sequence. However, after each rewind, the same dialogue between Chloe and Nathan repeats over and over again until the in-game task is solved. From a psycholinguistic perspective, the repetitive nature of the time reversal mechanics may result in better recall of dialogue used during action sequences.

2.9.4 Narrative choices in LIS

Another key feature of the time reversal mechanics which results in the repetition of in-game dialogue occurs when the game allows players to remake specific in-game choices. These choices will affect how the narrative of the game unfolds and result in unique game endings. Players may therefore be incentivized to redo these in-game narrative choices in order to accomplish different outcomes later on in the game. One example of the narrative choice mechanics described above occurs after Max saves Chloe from being shot in the girls' restroom by Nathan. As Max pulls the fire alarm in order to save Chloe, the entire school is evacuated. Before Max exits the school building, she must talk with principal Wells. During the exchange, principal Wells notices that Max is distraught and asks what is bothering her, at which point the player is given one of two choices: 1) to report that Nathan threatened Chloe with a gun or 2) to hide the truth from principal Wells. The narrative path of the game will differ depending on which narrative choice the player makes, a fact which the game actively draws attention to.

Should the player choose to report Nathan, Max expresses doubt that the principal will be able to do anything due to the influence that Nathan's family exercises over the school and offers the player the opportunity to reverse time and change the story. Should the player choose to hide the truth, Max again expresses discomfort in the decision citing that principal Wells suspects that Max is lying and that she may have jeopardized her scholarship. Just as with the first choice, Max suggests that the player could reverse time and tell principal Wells the truth. Regardless of which option is chosen, the game encourages careful consideration of each choice by informing the player that the choice will carry consequences later on in the game. Players may therefore be incentivized to utilize the time reversal mechanics to redo such narrative choice sequences. Using the time reversal mechanics to remake a specific choice in the game will expose the player to the same dialogue leading up to the choices, the repetition of which will arguably enhance recall of the TL vocabulary relevant to the dialogue used during these exchanges.

2.9.5 Spaced gaming in LIS

The comparatively low levels of linguistic down-regulation, as may be seen with the action sequences described above, coupled with the repetitive nature of the time reversal mechanics may arguably increase the recall of TL vocabulary acquired by language learner players. However, the lack of a battle system in LIS, and the process of in-game grinding associated therewith, may work to decrease the spacing intervals between the presentation of in-game vocabulary relative to other digital RPGs. Although the high levels of linguistic up-regulation embedded within the gameplay mechanics of LIS may confer benefits to the acquisition and retention of TL vocabulary, the absence of in-built spacing intervals derived from linguistic down-regulation may yield comparatively lower rates of long-term retention. Lower frequencies of linguistic down-regulation may function to decrease the efficacy of long-term retention of TL vocabulary acquired during gameplay by decreasing the total gameplay interval. The longer a player remains engaged with gameplay mechanics conducive to vocabulary acquisition and retention, the greater the potential for encoding variability. Furthermore, the more frequent the oscillations between in-game linguistic up- and down-regulation, the greater the potential for the spacing effect to enhance long-term retention of TL vocabulary acquired during gameplay.

Although LIS does not incorporate a linguistically down-regulated battle system which may result in intervals of grinding to achieve experience points or in-game money, it does incorporate an in-game activity which also functions to increase the total gameplay time common to most digital RPGs, namely side quests. Side quests are optional in-game objectives, the completion of which has no bearing on the progression through the main narrative of the game (Machado, Santos, & Dias, 2017). Players of digital RPGs may be incentivized to complete side quests to collect rewards such as rare in-game items or to garner a deeper insight into the game's narrative. The side quests in LIS play out in much the same fashion as with other digital RPGs. One example of a side quest in LIS is the photo collection side quest, for which the player is provided clues to specific events or locations in the game. Once the event or location has been identified, the player photographs it which unlocks an in-game "achievement." Fulfilling these achievements has no bearing on progressing through the game other than to collect the in-game photos. These side quests, superfluous as they may be, nevertheless serve as a critical component to increasing the spacing intervals between the presentation of the TL vocabulary used during required in-game sequences.

2.9.6 Information consolidation in LIS

The lack of linguistically down-regulated gameplay mechanics in LIS which elicit gaming styles conducive to time-intensive grinding (e.g., continuously battling enemies to gain experience points and/or in-game money) may significantly reduce the total gameplay time and interval. Total gameplay time (i.e., how many hours it takes to complete the game) is directly associated with the total gameplay interval (i.e., how long the interval of engagement with the game lasts in real-world time). It can therefore be assumed that digital games with longer gameplay times will take up longer intervals of real-world time relative to digital games with shorter total gameplay times. In other words, the interval of engagement with a digital game that takes an average of 20 hours to complete may be measured in weeks, whereas a digital game that takes an average of 100 hours to complete may be measured in months or even years. As the length of engagement affects memory consolidation of to-be-learned material, both the total gameplay time and interval of digital games may be indicative of their potential to elicit long-term retention of TL vocabulary. Studies looking at TL vocabulary acquisition using digital games suggest that longer periods of engagement, from two to six months, may achieve higher levels of vocabulary consolidation (Cobb & Horst, 2011; Wrobetz, in press). It therefore stands to reason that digital games with longer average gameplay times are more conducive to long-term consolidation of TL vocabulary acquired through gameplay.

LIS has a short gameplay time relative to other commercial single-player RPGs. With an average total gameplay time of approximately 16 hours (HowLongToBeat, 2021), LIS is considerably shorter than other single-player RPGs produced by Square Enix such as *Final Fantasy IX* with an average total gameplay time of approximately 51 hours (HowLongToBeat, 2021). These average total gameplay times may also change considerably when factoring in the amount of optional gameplay mechanics available (e.g., side quests) and differing gaming styles. The gaming style known as "completionist" is a term referring to players who attempt to complete every optional side quest available in the game. Average total gameplay times for LIS completionists top out at approximately 19 hours (HowLongToBeat, 2021), whereas *Final Fantasy IX* completionists average approximately 84 hours of total gameplay time (HowLongToBeat, 2021). These large differences in average total gameplay times indicate that more time-intensive single-player RPGs such as *Final Fantasy IX* may be better able to capitalize on the spacing effect relative to LIS. Therefore, from a psycholinguistic perspective, TL vocabulary acquired while playing LIS may undergo less long-term consolidation relative to more time-intensive single-player RPGs due to comparatively shorter total gameplay intervals. However, the narrative choice mechanics of LIS (see Table **2.5**) may increase the game's replayability and incentivize players to increase the total gameplay interval.

2.9.7 Multiple endings to increase long-term engagement

The choice-dependent narrative structure in LIS gives rise not only to different narrative paths through the game but also to unique endings. Officially there are two possible endings in LIS (GameGuides, 2021). However, there are a total of 62 in-game choices, most of them optional, which alter the content and progression of the game's narrative structure (Life Is Strange Wiki, 2021). For example, the narrative choice of making fun of or comforting Victoria after she has paint splashed on her will either result in Victoria leaving an unflattering picture of Max up on social media if the player chooses to make fun of Victoria or in Victoria deleting said social media post if the player chooses to comfort Victoria. When the player accesses Victoria's laptop to search for evidence that she is in fact the real source of conflict between Dana and Juliet, the choice made by the player will change the dialogue of the game by way of Max either commenting on the fact that Victoria deleted the photo or that the post still remains. Moreover, if the player chooses to make fun of Victoria, Max takes an instant photo of the paint-covered Victoria, which will prompt unique dialogue exchanges with both Warren and Chloe.

Players who are motivated by the narrative choice mechanics and how they affect the unfolding of the game's narrative structure may be incentivized to replay the game in order to experiment with different narrative outcomes elicited by the narrative choice mechanics. The overall replayability of LIS is also arguably greater than in more time-intensive single-player RPGs such as *Final Fantasy IX* precisely due to the shorter play-through time, which allows players to experiment with the different versions of the narrative structure without the considerable time commitments that would be required from longer games. Language learner players of LIS who are motivated by the game's replayability may therefore experience more long-term consolidation of TL vocabulary acquired during gameplay due to longer total gameplay intervals relative to players who only play through the game once.

2.9.8 Multimodal presentation in LIS

The in-game environment in LIS is permeated with interactive multimodal vocabulary, the use of which may enhance TL vocabulary acquisition/retention. Not only is the majority of linguistic in-game content voice acted and accompanied with subtitles, but the interactive gameplay mechanics provide players with multimodal linguistic descriptions of large portions of the in-game environment. As the player navigates the in-game environment, labels of objects and NPCs that can be interacted with will appear. Upon approaching a labeled object, it becomes highlighted and an action which the player may perform on the object appears. Max provides commentary regarding her feelings about each object through an inner monologue. The inner monologue commentary on in-game objects is also voice acted and accompanied with subtitles. Unlike most digital RPGs, the environmental label mechanics in LIS provide a continual stream of multimodal linguistic input (i.e., simultaneous visual and textual representation of in-game objects) which may be interacted with to provide an additional audial and contextual component to a large portion of the in-game environment. From a psycholinguistic perspective, the sheer volume of multimodal linguistic input may theoretically enhance TL vocabulary acquisition and recall when played in SLA contexts.

2.9.9 Guided play in LIS

The similarity of the gameplay mechanics in LIS to the fundamental principles of guided play may also theoretically enhance TL vocabulary acquisition/retention. Apart from the similarities to guided play shared with other digital RPGs (i.e., corrective feedback mechanics coupled with scaffolded semi-freeform play), the script choice mechanics in LIS function similarly to scaffolded social exchanges during guided play sessions. During social exchanges in LIS, the player is given the power to control the direction of the conversation through the script choice mechanics. The content of the in-game conversations, however, is ultimately based on scripted dialogue, and therefore the only level of control that the player may exert over these conversations is which portions of in-game dialogue are uncovered by the script choices made during social exchanges. It also goes without saying that the player need not actually produce any language in order to engage in these in-game social exchanges. During guided play sessions, the context of play is pre-determined, but ultimately the child directs the course of these exchanges by choosing what to talk about and when. Similarly, in LIS, the context of the game is also pre-determined through the game's scripted narrative structure, but the player can choose the direction, content, and duration of a large portion of social exchanges.

A language learner playing LIS who is particularly motivated by the game's role-playing narrative elements may be incentivized to use the predetermined context in order to comprehend which script choice will allow the player to accomplish a desired in-game task or unlock a desired narrative path. The semi-freeform structure of these in-game social exchanges resembles how a conversation may play out during a guided play session during which the child uses a predetermined context to engage in social exchanges for the purpose of accomplishing a desired task. Furthermore, the element of active role playing, whether through identity projection onto inanimate objects such as toys or through the mimicry of different social roles (e.g., when playing "house"), is common among children at play (Bowers & London, 1965) and makes up one of the principle ludic elements of RPGs, whether digital or otherwise (Montola, 2007). As guided play, or child-directed play in a predetermined context, has been shown to enhance vocabulary acquisition, then the player-directed script choice mechanics within the predetermined context of the narrative structure in LIS may, on a theoretical basis, be conducive to TL vocabulary acquisition in SLA contexts.

2.9.10 Summary of memory mechanisms in LIS

When compared to other genres of single-player RPGs, particularly those with a central focus on linguistically down-regulated gameplay mechanics such as battle systems, LIS may be particularly effective at eliciting TL vocabulary acquisition and long-term retention. Not only does LIS incorporate a high volume of linguistically up-regulated gameplay with a central focus on the comprehension of in-game tasks that use the script choice mechanics but other gameplay features such as the time reversal and narrative choice mechanics may function to increase a language learner player's exposure to processes beneficial to TL vocabulary acquisition and retention. In particular, the gameplay mechanics in LIS demonstrate the ability to theoretically enhance TL vocabulary acquisition/retention by eliciting inferential processing, repetition, spaced gaming, long-term consolidation, multimodal TL input, and a player-directed narrative structure similar to the principles of guided play (see Table **2**.**6**). Furthermore, the script/narrative choice mechanics which allow the player to direct the course of the narrative structure allow for quasi language production within the context of the game, an SLA-beneficial feature common in MMORPGs through the social gameplay mechanics but which is wholly absent from the grand majority of single-player RPGs. Therefore, from both a sociocultural perspective and a psycholinguistic perspective, engagement with LIS in SLA contexts may benefit the acquisition and retention of TL vocabulary.

Table 2.6 Memory mechanisms in LIS gameplay mechanics

Memory mechanism	LIS gameplay mechanics
Inferential processing	In-game tasks/objectives with linguistic input acting as a source for inferential processing allows the player to confirm/correct inferences of unknown TL vocabulary through explicit corrective mechanics (e.g., forced time reversal), non-explicit corrective mechanics (e.g., game-provided clues), and negotiation with an interactive environment (e.g., accessing additional linguistic input through item descriptions).
Repetition	Certain in-game tasks utilizing script choice mechanics require the utilization of information from social exchanges. Once relevant information has been acquired by the player, the time reversal mechanics may be used to redo the conversation and provide correct script choices to move the game forward, thus repeatedly exposing the player to TL input. TL input during timed action sequences (i.e., linguistically down-regulated sequences focusing on in-game controls) will be repeated every time the player uses the time reversal mechanics. Players who are motivated by the game's multi-path narrative structure may be incentivized to redo narrative choice sequences in order to observe their effects on the unfolding of the game's narrative, thus repeating any TL input used therein.
Spaced presentation	The length of time between required in-game sequences may be lengthened through the engagement with optional narrative side quests such as: in-game achievements (e.g., the photo collection side quest) and optional narrative choice sequences.
Consolidation	The multiple-path narrative structure and multiple endings may incentivize players to replay the game, thereby functionally increasing the total gameplay interval.
Multimodal presentation	The in-game environment presents the player with multimodal TL input with voice-acted/subtitled narrative content and labeled interactive items with visual, audial, and textual descriptions.
Guided play	The script choice mechanics resemble social exchanges in the guided play methodology with player-directed script choices in optional sequences (choosing how much information is needed), required script choices (choosing your own narrative paths), and the role-playing element in script choices (choosing how you want the character Max to behave).

2.10 Research questions

With the potential, as outlined above, of the smartphone application of LIS to elicit TL vocabulary acquisition and retention within formal foreign language educational contexts, this study investigates four research questions (RQs) to investigate the efficacy of such a learning intervention in EFL courses at a private Japanese university. The RQs were used to structure the methodology and data analysis procedures used in the present study and will be discussed respectively in the following chapters. Research question 1: How effective is the LIS gaming intervention at eliciting TL vocabulary acquisition and retention when used in conjunction with pre-determined testing protocols relative to a control group? Research question 2: How do the gameplay mechanics (i.e., the rules that govern a players' actions within the context of the game) in the smartphone application of LIS affect vocabulary acquisition and retention of TL vocabulary? Research question 3: How will the LIS gaming intervention correlate to students' perceptions of and attitudes towards English smartphone gaming? Research question 4: How will the depictions of US culture in the narrative content of LIS affect how the students inform their impressions on US culture?

Chapter 3 Methodology

As outlined previously, the game LIS, a choose-your-own-adventure, single-player RPG, was chosen as the focus of this study due to its incorporation of gameplay mechanics which may be particularly conducive to SLA (see Table 2.1 and Table 2.6). The smartphone edition was chosen as the gaming platform due to the affordances of the platform (e.g., mobility and scalability) in formal academic environments (see Table 2.2). The curriculum designed around the incorporation of LIS in university-level EFL courses attempted to address the research questions presented in chapter two. After briefly introducing the basic properties and mechanics of LIS, this chapter will outline how the study was conducted.

3.1 Game overview

The version of LIS used in the present study was the free-to-play smartphone edition of the first chapter *Chrysalis* (Square Enix, 2015). All study participants were required to download the application to their personal smartphones and use personal headphones during in-class gaming sessions.

3.1.1 Language selection for in-game subtitles

As the smartphone edition of LIS was implemented in university-level EFL courses, study participants were not only required to play the game in English but also to turn on the English subtitles. As the use and language setting of in-game subtitles is entirely optional within the gameplay mechanics, study participants were explicitly shown how to turn on the English subtitles with an in-class demonstration.

3.1.2 Setting of LIS

In the game LIS, the player controls the main character Max, a girl enrolled in a high school known as Blackwell Academy in the Pacific Northwest in the USA. The player can interact with most objects embedded in the in-game environment which will elicit scripted responses from the main character. The player can also interact with most NPCs embedded in the in-game environment, interactions which will occasionally result in the player being given a choice affecting how the game's narrative structure unfolds. The player has the ability to reverse time in order to undo/redo choices and in-game event sequences. These player-directed choices which result in multiple narrative paths through the game coupled with the time reversal mechanics and a highly interactive in-game environment which continually exposes players to TL input were identified as desirable gameplay mechanics for use in SLA contexts.

3.2 Institution and course context

This study was conducted in three EFL courses at a university in Japan. The three courses in which this study was conducted were Business English, Communication English, and English Reading. The curricula of all courses were designed and implemented by the author of this study to improve general EFL skills within the context of scenarios which may require business-oriented English proficiency. The LIS study was implemented in the above courses as one component among other course-specific activities which included textbook-oriented courses of study, extensive reading, and oral presentations. Participation in this study therefore made up a certain percentage of the enrolled students' grades. However, each student's grade within the LIS curriculum was determined by the timely completion of study-specific materials and not by individual performance on said materials. All enrolled students were informed of the study's grading scheme and willingly participated in the study in exchange for course credit on the condition of anonymity upon data publication (see Appendix A). Furthermore, the study's grading scheme was designed in such a way so as to ensure that non-compliance with the study's testing schedule would not result in a failing grade.

3.3 Student participation details

3.3.1 English and Japanese proficiency

There was a wide range of English proficiency levels among the participants in this study. The students' English proficiency level could vary, as well as could be ascertained from the author's evaluation based on the quality of student course work and in-class exchanges with the students, from novice-low all the way to advanced-high using the American Council on the Teaching of Foreign Languages (ACTFL) proficiency guidelines (Breiner-Sanders, Lowe, Miles, & Swender, 2000). Some of the disparities in English proficiency level may have been due to age differences/level of exposure to EFL curricula as participants in this study ranged from being first-year to fourth-year students. As a point of observation, students' level of motivation to participate in class was a better predictor of English proficiency than grade level. This observation is supported by research examining motivation and EFL proficiency (Bernaus & Gardner, 2008; Weda, 2018). Other potential causes of the wide range of English proficiencies were the socioeconomic and cultural backgrounds of the enrolled students (Hill & Giammatteo, 1963; Nikolov & Csapó, 2018; Weda, 2018). With specific regard to the cultural differences between the participants, at least three participants were known to be exchange students. All other participants were assumed to be native Japanese speakers regardless of their ethnic/cultural backgrounds (e.g., immigrants to Japan). As a specific score on a standardized English proficiency examination such as TOEIC was not a prerequisite for enrollment, all quantitative measurements of English proficiency are based on study-specific materials created by the author.

As outlined above, three students who participated in this study were exchange students. Two were from China and one was from South Korea. Although the game was played in English and the overarching purpose of this study was to measure the efficacy of LIS to serve as an instructional medium in formal EFL classroom environments, due to the overwhelming majority of native Japanese speakers, a certain amount of Japanese was used throughout the study. Therefore, a high level of Japanese proficiency was required in order to successfully participate in this study. All three exchange students had received the N1 qualification (i.e., the highest rank) on the Japanese Language Proficiency Test (JLPT). Furthermore, as the students' level of Japanese proficiency was not hindering their participation in other courses at the university, the level of Japanese proficiency demonstrated by the three exchange student participants was determined to be suitable for the inclusion of their data in the final analysis.

3.3.2 Study participation and data collection procedures

A total of 172 students willingly participated in this study. As participation guaranteed anonymity, each participant's name was assigned an alphanumeric value indicating the study group and participant identification number. The "PS" study group denotes the pilot study group. A total of 30 students participated in the PS group. The "LG" study group denotes the *Life Is Strange* gaming group (i.e., the treatment group), and the "LC" study group denotes the *Life Is Strange* control group (i.e., the non-treatment group). A total of 99 students participated in the LG group, and 43 students participated in the LC group. The difference in study group size was primarily the result of LG group participants being enrolled in courses in which the LC treatment was conducted. As participation in the LG group was prioritized, the LC group size was negatively affected. Further, as a result of the Covid-19 pandemic, enrollment in the EFL courses in which this study was conducted diminished significantly, reducing the total number of LC group participants even further. Although equal sample sizes would be preferable, the total LC group sample size is over 50% of the total treatment, which is acceptable for statistical analysis (White, 2019).

Regardless of the study group, each participant was presented with the same data publication consent statement upon taking each study-related test and/or survey (see Appendix A). None of the student participants voiced any objections to the anonymous use of their coursework data in the present study. Although participation in this study comprised a certain percentage of the student participants' final course grades, in the event that a student participant voiced an objection to the use of his/her data in this study on ethical grounds, alternative coursework would have been provided in lieu of participation in the present study.

3.3.3 Pilot study group participation parameters

The pilot study took place in the spring semester of 2019 (from April 2019 to July 2019) in one EFL course. Although the PS group played through the same portions of LIS as the LG group, the PS group did not take any of the vocabulary tests, comprehension tests, or surveys that were taken by the LG and LC groups. The purpose of the PS group was to ascertain how much time would be necessary to complete the LG study program and to help design the tests and surveys that would be used later on in the study program by both LG and LC groups. Participation in the PS group therefore only required completing the game in class. While the PS group played the game, the author monitored how long the game took to complete and which portions of the game were particularly difficult to complete. This information was subsequently utilized to create the study's gaming and testing schedule used for the LG and LC groups. At the end of the semester, the students in the PS group were required to complete group oral presentations, one of the topics of which was about playing LIS in class. A total of 10 students chose to present on playing LIS, the presentation materials of which are presented in Appendix B and Appendix C. The following discussion of testing procedures only refers to the LG and LC groups, as the PS group did not take any tests in this study.

3.3.4 LG and LC group participation parameters

In order to qualify for inclusion in the data analysis portion of this study, a number of conditions had to be met. First and foremost, as the data analysis of the vocabulary and comprehension test data was time sensitive, all vocabulary tests, comprehension tests, and surveys had to be submitted on time. The only exception to the test/survey submission deadlines was for the one-semester delay vocabulary test, any submissions of which were accepted after a minimum of 13 weeks had passed since the submission of the vocabulary pre-test. Second, all participants in the LG group had to play LIS with English audio and English subtitles throughout the entire study period. Third, all participants had to refrain from using dictionaries both during gameplay and during testing. If any of the above conditions were not met, the participant's data were removed from the final data analysis.

3.3.5 Switch to online learning during the Covid-19 pandemic

In the fall semester of 2019 (from September 2019 to January 2020), all courses were conducted in the classroom. Therefore, all students were required to be physically present in the classroom to participate in the study, during which time the author physically monitored the participants' in-game language settings and adherence to the study participation parameters regarding the use of dictionaries. Due to the influence of the Covid-19 pandemic, the instruction method of the courses shifted to online learning during the spring semester of 2020 (from April 2020 to July 2020) and then to a hybrid instructional method (i.e., simultaneous in-person and online learning) in the fall semester of 2020 (from September 2020 to January 2021) and the spring semester of 2021 (from April 2021 to July 2021). Due to the online nature of the instructional methods during the Covid-19 pandemic, student's in-game language settings and dictionary use during testing could not be physically monitored.

To ensure that students were adhering to the study parameters during semesters with online courses, all participants were required to submit screenshots to provide proof that they had completed the assigned section of the game for that day's class and that the game's language settings were set to English. Along these lines, students who were absent from class were still allowed to participate in the study provided that they sent the author screenshots proving the same parameters described above. Students in the LC group, as they did not play LIS and only took the vocabulary tests, were not required to submit screenshots and all timely submissions of the vocabulary tests were included for data analysis in this study regardless of whether the participant was present in class. In replacement of the physical monitoring of the participants' dictionary use during testing, a test question was added into all vocabulary tests requiring students to agree to the compliance of the no-dictionary rule. The question was posed in Japanese, "次の質問を答えるとき、辞書を使わないことを誓います (I hereby pledge to not use a dictionary in the answering of the following questions)" to which the participants could either answer, "誓います (I do)" or "誓いません (I do not)." The English translations listed above are included for clarification purposes only and were not used during formal testing.

3.3.6 Instances of non-compliance and data removal

In total, 42 participants' data were removed, 20 from the PS group, 21 from the LG group, and 1 from the LC group (see Table 3.1). None of the participants were physically observed using dictionaries during testing, however there were two instances of study participants not pledging to abide by the no-dictionary rule on the LG group's vocabulary post-test in the spring semester of 2020. Upon clarifying both students' responses to the dictionary compliance question, both students claimed to have accidentally selected the "do not agree" response. Furthermore, both students claimed to have completed the tests without the use of dictionaries. In reviewing both students' responses to the dictionary compliance question on previous vocabulary tests, it was found that both students complied with the no-dictionary rule on all other vocabulary tests, thereby lending credence to their claims of accidental selection. All data from both participants were therefore included for final data analysis. The datasets after the removal of non-compliant/non-relevant data are hereinafter referred to as modified study groups.

Table 3.1 Description of study groups

Group name and size		Instances of removal	Reasons for removal		Group size after data removal
PS group	30	20	No relevant data to analyze	20	10
LG group	99	21	Non-submission of test(s)	18	78
			Changing game language settings	3	
LC group	43	1	Non-submission of test(s)	1	42
Total	172	42		42	130

3.3.7 Instances of repeat participation

Within the modified LG group, there were a total of eight participants who completed the study twice. Three of the 11 repeat participants' data from the spring semester of 2020 were removed due to non-submission of tests or non-compliance with the study participation parameters. The three participants mentioned above, however, were all able to successfully complete the study in the fall semester of 2019, and their data is therefore included in the final data analysis. In addition, one repeat student's data were removed from the fall semester of 2019 due to the non-submission of tests but were left in for the spring semester of 2020 due to successful completion of the study program. None of the repeat students participated in the study twice in the same semester. Finally, of the eight students who successfully completed the study twice, these students were assigned two study identification numbers to correspond to the semester of participation. None of the participants in the LC group participated in the study more than once. No statistically significant skews of data from repeat participation were observed in the final data analysis (see Table 4.4), therefore all data from repeat participation were left in.

3.4 Study details and in-class procedures

3.4.1 Academic schedule

The present study was conducted over six semesters, or three academic years. The collection of data being presented in this study commenced in the spring semester of 2019 and concluded in the fall semester of 2021. As this study was conducted at a Japanese university, the semesters referred to in this study conform to the standard schedule of academic semesters at institutions of higher education in Japan. A standard semester of study in Japan typically consists of 15 weeks of study spanning three and a half months. An academic year in Japan commences from the spring semester, which typically runs from April to July, and concludes with the fall semester, which runs from September to January. Each of the courses in which the present study was conducted consisted of 15 one and a half hour classes meeting once weekly.

3.4.2 Designing the study schedule

The pilot study took place in the spring semester of 2019. As outlined previously, the purpose of the PS group was to help determine the length of the study program and the target vocabulary for the following semesters. The study schedule for the PS group therefore differs from the LG and LC group study programs and took place over four weeks. The PS group study schedule is summarized in Table 3.2. The results of the pilot study determined that the study program for the LG and LC groups would have to be carried out over five weeks, one week for study program introduction, three weeks for gaming and testing, and one week for study program conclusion.

Table 3.2 PS group study schedule

Study week	Description of in-class activity
Week 1	Students were informed of the gaming project and instructed to download the LIS application on their smartphones before the next week's class. Furthermore, students were instructed to bring personal headphones in order to listen to the voice-acted, in-game dialogue without disturbing other students.
Week 2	Before commencing gameplay, students were instructed on how to switch the game's English subtitles on. The entire 90 minutes of class was used to play through the game while the author monitored how the game was proceeding. Just before class ended, students were instructed to not play the game until next week's class.
Week 3	As the author noted that most students were over halfway through the chapter of LIS covered in the present study, students were instructed to keep playing the game from the point they left off in last week's class and that they would be dismissed from class once finished with the game. Roughly one third of the class finished the game before the end of class. The rest of the students required the full 90 minutes to complete the game.
Week 4	An in-class discussion was held to determine which parts of the game (vocabulary or otherwise) were difficult to comprehend and/or complete.

3.4.3 LG group study and testing schedule

The following sections will detail the LG group study schedule, summarized in Table 3.3. As the main goal of the LG group was to measure the efficacy of the LIS gaming intervention on EFL vocabulary acquisition and retention, the project introduction during the first week of the study program included the statement that students would be playing LIS in order to ascertain what impact the game had on the development of EFL skills such as vocabulary. It was also explained that since one of the study's goals was to measure the game's impact on the students' EFL vocabulary, no dictionaries would be allowed during gameplay or testing. The study program introduction for the LG group also introduced the study's grading scheme and testing schedule. After the introduction, the LG group was given a 45-question vocabulary pre-test (see Appendix D) to serve as an evaluation of the students' knowledge of the vocabulary that they would subsequently be tested on over the next four weeks. Upon completion of the vocabulary pre-test, no corrective feedback was provided to the students. Furthermore, an additional pre-game survey was administered after the completion of the vocabulary pre-test (see Appendix E). The pre-game survey was designed to gather both quantitative and qualitative data on the students' engagement with smartphone games, their perceptions on the efficacy of utilizing smartphone games as an instructional medium in EFL contexts, and the perceptions of the TL culture, namely US culture.

Table 3.3 LG group study schedule

Study week	Description of in-class activity	Tests/surveys taken after in-class activities
Week 1	Project introduction	Pre-game vocabulary test of all test items with no CF Pre-game survey
Week 2	Introduction to 17 vocabulary test items in GS1 Gaming session 1 (GS1)	Immediate vocabulary test 1 with CF Comprehension test A with CF
Week 3	In-class discussion about content from GS1 Introduction to 13 vocabulary test items in GS2 Gaming session 2 (GS2)	Immediate vocabulary test 2 with CF Comprehension test B with CF
Week 4	In-class discussion about content from GS2 Introduction to 15 vocabulary test items in GS3 Gaming session 3 (GS3)	Immediate vocabulary test 3 with CF Comprehension test C with CF
Week 5	In-class discussion about content from GS3	Post-game vocabulary test of all test items with CF Post-game survey
13+ weeks	Requests sent to participants to take the one-semester delay (OSD) vocabulary test (non-compulsory)	OSD vocabulary test of all test items with CF

Note. Corrective feedback (CF) refers to automatic grading in Google Forms.

3.4.4 In-class procedure before gaming sessions

The game itself was played over three weeks (i.e., on weeks two, three, and four of the study program). The three gaming sessions split the first chapter of LIS (i.e., the chapter covered in the study) into roughly equal segments of gameplay time measured during the PS group's gaming sessions. Before each gaming session, the LG group was introduced to each word and the sentence in which the word would appear during gameplay (i.e., the vocabulary test items). Before the gaming sessions on weeks three and four, in-class discussions were also carried out regarding the narrative content of the previous week's gaming sessions. However, no such content discussion took place on week two as there was no prior gaming session to discuss. After presenting the vocabulary test items and carrying out in-class content discussions of the previous weeks' gaming sessions, the students were instructed to play the game, pay attention to the in-game dialogue, and utilize the context of the game in order to figure out the meaning of the introduced vocabulary test items. Once students reached the predetermined check point for that week's gaming session, students were instructed to inform the author, after which the post-game testing procedure was initiated.

3.4.5 In-class procedure after gaming sessions

After each weekly gaming session, the LG group took vocabulary tests with vocabulary drawn from each respective gaming session (see Appendix F, Appendix G, and Appendix H). In addition to the vocabulary tests, the LG group also took three weekly comprehension tests with questions that corresponded to the content from each week's respective gaming session (see Appendix I, Appendix J, and Appendix K). All tests were conducted over Google Forms via hyperlinks posted to a university-affiliated website. After the completion of each vocabulary test and comprehension test, the students were automatically provided with corrective feedback (i.e., their total test score, which questions they answered incorrectly, and the correct answers to each question) for all vocabulary and comprehension tests via Google Forms.

3.4.6 LC group study and testing schedule

The following sections will detail the LG group study schedule, summarized in Table 3.4. The LC group, who did not play the game, took only the vocabulary tests at the beginning of class and then continued with their normal course of study not related to the game. In contrast to the LG group vocabulary treatment, the vocabulary test items and their accompanying sentences were not introduced to the LC group prior to testing. The vocabulary test items and their accompanying sentences were not introduced to the LC group due to the fact that there was no treatment to complete before the actual test and the test items on the LG vocabulary tests and LC vocabulary tests are identical. A testing procedure which included the presentation of the vocabulary test items and their accompanying sentences followed immediately by the exact same test questions was determined to be redundant, and that procedure was therefore omitted from the LC vocabulary test treatment. Just as with the completion of the LG vocabulary tests, the LC group received the same corrective feedback as the LG group. Since a normal lecture was scheduled to take place after the vocabulary tests in the LC group courses, a buffer of five minutes starting from the submission of all vocabulary tests for the students present in class was utilized in order to ensure that all students had a minimum of five minutes to review the corrective feedback. After the five-minute buffer had elapsed, each respective LC group lecture resumed as normal.

Table 3.4 LC group study schedule

Study week	Description of in-class activity	Tests/surveys taken after completion of in-class activities
Week 1	Project introduction	Pre-game vocabulary test of all test items with no CF
Week 2	No testing-related in-class activity	Immediate vocabulary test 1 with CF
Week 3	No testing-related in-class activity	Immediate vocabulary test 2 with CF
Week 4	No testing-related in-class activity	Immediate vocabulary test 3 with CF
Week 5	No testing-related in-class activity	Post-game vocabulary test of all test items with CF
13+ weeks	Requests sent to participants to take the one-semester delay (OSD) vocabulary test (non-compulsory)	OSD vocabulary test of all test items with CF

Note. Corrective feedback (CF) refers to automatic grading in Google Forms.

3.4.7 Post-test procedure for LG and LC groups

The results from the vocabulary tests taken on weeks two, three, and four of the study program in both LG and LC groups are referred to as the immediate test results due to the fact that the test results reflect the immediate effect of the treatment that each group received (i.e., game vs. no game) without the benefit of corrective feedback. On the fifth and final week in the study program (see Table 3.3 and Table 3.4), both LG and LC groups took a 45-question vocabulary post-test at the beginning of class (see Appendix D). This vocabulary post-test featured the same vocabulary test items and their accompanying sentences that appeared in the vocabulary pre-test and vocabulary tests one, two, and three from the three weeks of treatment in weeks two, three, and four of the study program. The vocabulary post-test therefore has an in-built three-week delay for the vocabulary test items from vocabulary test one (see Appendix F), two-week delay for the test items from vocabulary test two (see Appendix G), and a one-week delay for the test items from vocabulary test three (see Appendix H). Just as with vocabulary tests one, two, and three, the vocabulary post-test provided the students with corrective feedback via a Google Forms webpage link. The corrective feedback review buffer provided for the LC group after the immediate vocabulary tests was increased to 10 minutes for the vocabulary post-test due to the increased test question load. In addition to the vocabulary post-test, the LG group was also required to take a post-game survey (see Appendix L) to measure shifts in the same quantitative and qualitative data gathered in the pre-game survey (see Appendix E).

The final vocabulary test issued was a voluntary one-semester delay vocabulary post-test (OSD post-test) for both LG and LC groups. Unlike the vocabulary tests, comprehension tests, and surveys conducted during the five-week study programs, the participation in the OSD post-test had no effect on the final grade of the student participants. The OSD post-test, like the vocabulary pre- and post-tests, consisted of the same 45 questions (see Appendix D) and provided corrective feedback upon submission of the test via Google Forms. A request to take the OSD post-test was sent out to all participants who had completed the vocabulary post-test at least 13 weeks prior. As the submission of the OSD post-test was entirely voluntary, data from students who chose not to take the OSD post-test, but who had completed the mandatory vocabulary tests from the five-week study program, were included for final analysis.

3.4.8 Supplementary in-class activities for the LG group

As outlined above, the LC group's participation in the study program (see Table 3.4) was limited to the submission of the vocabulary tests during the five-week study program and the voluntary submission of the OSD post-test. After the LC group had submitted the vocabulary tests, the student participants received a standard lecture for the course in which they were enrolled which did not incorporate any content about LIS. The LG group, however, had to engage with the single-player RPG in the formal EFL context of the classroom as a learning intervention. The courses in which participants in the LG group were enrolled therefore incorporated in-class activities which utilized the game's content to carry out the pedagogical goals of each specific course. The supplementary out-of-game activities took the form of in-class discussions and oral presentations about the narrative content of LIS. The content of both the in-class discussions and oral presentations were on a staggered schedule relative to gameplay so as to never provide hints about the answers to vocabulary and comprehension test items. The oral presentations were scheduled to take place after the five-week study program had concluded, and the in-class discussions always revolved around the gameplay content from the previous week (see Table 3.3). Participation in these in-class discussions and oral presentations was required and affected each participant's course grade. Although no student participant raised an objection about participating in the present study, the author had prepared alternative methods of participation in such an event.

3.4.9 In-class narrative content discussions

The in-class discussions regarding game content took place in the three EFL courses in which the present study was conducted. There were two categories of discussion questions utilized in this study: open-ended discussion questions and content review questions. The topics of open-ended questions varied depending on the gaming session and semester but generally required students to summarize the in-game content from the prior week's gaming session, discuss how well they understood said content, and analyze how US culture is portrayed in the game (see Appendix M). Each student was required to prepare responses to the open-ended questions in English and share their responses with student discussion groups first and again later on in full-class discussions. The content review questions did not require the students to prepare individual responses, however the students were often called on to try and recall specific points of discussion before receiving an explanation from the author. Examples of content review questions included asking students who specific characters are, what specific elements of the story are about, and where the story takes place (see Appendix M). After calling on a number of students to answer the content review questions, a detailed explanation was then given by the author. These in-class discussions about the prior week's game content always preceded the gaming sessions (see Table 3.3).

3.4.10 LIS oral presentations in the PS group

The other compulsory in-class project required by some of the participants in this study took the form of oral presentations about playing LIS in EFL contexts. In contrast to the in-class discussions, which took place in all of the courses included in this study, the LIS oral presentations only took place in the course English Reading for the PS group and the course Communication English for the LG group. In the spring semester of 2019, the PS group was required to participate in group presentations on one of three topics covered in class: news article content review, advertising, or the portrayal of US culture in video games (i.e., LIS) (see Appendix N). Students could sign up for any presentation they wanted to but were limited to groups of four to five students. Each presentation topic was limited to two group presentations, therefore there were two group presentations on playing LIS that took place in the spring semester of 2019. Students were required to make a PowerPoint presentation (see

Appendix B and Appendix C) to answer a total of five questions about their impressions of US culture before and after playing LIS (see Table 3.5).

Table 3.5 Presentation questions for LIS oral presentation in PS group

Presentation topic 3: US culture in video games		
Presentation instructions	Please use the game *Life Is Strange* (Square Enix, 2015) to discuss how US culture is portrayed in video games.	
Presentation objectives	Your presentation should review how US culture is portrayed in the game *Life Is Strange*. Furthermore, your presentation should talk about the merits/demerits of using video games to portray culture.	
Presentation questions	1	What was your impression of US culture before playing?
	2	What was your impression of US culture after playing?
	3	How did your impression of US culture change afterwards?
	4	What are the merits of using video games to portray culture?
	5	What are the demerits of using video games?

3.4.11 LIS oral presentations in the LG group

In the fall semester of 2020, students enrolled in Communication English were also required to give oral presentations, but the oral presentation parameters differed in a number of ways from the group oral presentations that took place in the PS group. Due to comparatively lower student enrollment, the oral presentations were conducted individually instead of as a group. Furthermore, instead of allowing the students to choose their presentation topic from a limited selection of topics, the students were allowed to choose the questions they wished to answer from a limited selection of questions about the content of the semester (see Table 3.6). Furthermore, the students were required to compare three games and their applications in the EFL classroom that were played throughout the semester within the presentation, namely *One Night Ultimate Werewolf* (Bézier Games, 2013), *Keep Talking and Nobody Explodes* (Steel Crate Games, 2015), and LIS. The presentations of the two LG group students enrolled in the Communication English course are presented in Appendix O and Appendix P.

Table 3.6 Presentation questions for LIS oral presentation in LG group

Presentation topic: Using games to learn English	
Presentation instructions	You will be giving a presentation on your opinion about using games to learn English. You *must* answer the question about *Life Is Strange*. Then, answer at least *four* of the *nine* questions about the three games we played this semester (*One Night Werewolf*, *Keep Talking and Nobody Explodes*, and *Life Is Strange*) in a PowerPoint presentation. If you have any questions about this presentation or if you would like me to correct your English, please contact me *before* your presentation.
Mandatory question	How effective do you think the game *Life Is Strange* is at improving your English vocabulary? What other kinds of English language skills do you think it improves (if any)? What aspects of US culture did you learn when playing the game?
Optional questions	1 Considering the games that we played this semester, which game did you enjoy the most? Which game do you think helped improve your English the most? Why? 2 After having played through three different types of games in English, has your opinion of gaming in English changed? Do you think you are more likely or less likely to play games in English? 3 Do you think that you learn better through games or through more traditional educational materials (news articles, presentations, textbooks, etc.)? How are both learning methods similar? How are they different? 4 After having played through three different types of games in class, has your opinion of using games in class changed? Are you more likely or less likely to want to use games in class to learn English? 5 There are many genres of games. Considering the wide variety of games, what other types of games (outside of the three games we played in class) do you think would be useful to learn English? 6 Do you think that games will be used more often or less often in the future to learn English in educational settings? Why or why not? 7 What other kinds of subjects do you think games would be useful for? Why? 8 Before participating in this class, had you ever played a game in English before? What kind of game was it, and do you think it helped improve your English? 9 There are many different learning styles (e.g., visual, audial, kinesthetic, etc.). What kind of learner are you? Do you think games are beneficial or detrimental for a classroom with many different types of learners?

3.5 Vocabulary test, comprehension test, and survey structure

3.5.1 Word selection criteria

There were a total of 45 vocabulary test items selected for analysis (see Appendix D). The types of words and expressions chosen to be used as vocabulary test items included nonstandard English expressions (e.g., item number 5 “got hella cash”), words with multiple contextual applications (e.g., item number 19 “booty”), abbreviations and nonstandard spellings (e.g., item number 21 “photo op” and item number 17 “bizarro”), and modern English colloquialisms (e.g., item number 10 “selfie”). Due to the high volume of optional gameplay sequences available in LIS, all vocabulary test items were selected from mandatory gameplay sequences. The order of the vocabulary test items was randomized from the order in which they appear in the game.

3.5.2 Vocabulary test structure

All vocabulary test questions featured the vocabulary test items embedded in the sentence in which they appeared in the game (removed from the appendices), demarcated with parentheses. The participants were provided with the Japanese language instructions "答えるために括弧に入っている英単語を正しく日本語に変えて下さい。(To answer the questions, correctly change the English word in parentheses into Japanese)." All test questions were multiple choice and provided participants with four possible Japanese translations of the vocabulary test item. The multiple-choice selections featured one correct answer, one incorrect answer selected from the list of the other 44 vocabulary test items, one strong distractor, and one weak distractor. The multiple-choice selections for each vocabulary test item remained the same throughout all six tests, however the order of the multiple-choice selections were automatically randomized by Google Forms for every test generated.

All test questions had to be answered before submission on Google Forms. The multiple-choice format of the test with the required response settings was implemented in order to avoid non-responses from participants. Due to the grading structure of the study program, whereby the students' grades were determined by the submission of the relevant tests and not on the students' performance on the test, other test question response methods may have resulted in less accurate measurements of vocabulary test item comprehension. Should the test settings not have required each test question to be answered, participants could have submitted the test without looking at the test questions and still have received course credit for participating in the study program. Similarly, should the test response method have been open-ended instead of multiple choice, students could have answered "I don't know" for every single response without ever reading through the test questions and still have received course credit. Although the possibility of students randomly selecting multiple-choice responses for each test question still exists with the multiple-choice format used in this study's vocabulary tests, assuming the average of student scores remains above the statistical threshold of randomly selecting one of four multiple-choice responses on each test question (i.e., 25%), then it may be determined that the study's participants on the whole were not randomly guessing. Furthermore, relevant literature points to the benefits of multiple-choice testing on information retention (Little & Ligon Bjork, 2012).

3.5.3 Comprehension test structure

In addition to the vocabulary tests, the LG group was also required to take comprehension tests. The comprehension tests were designed to measure both how well the LG group understood the narrative content of each of the three gaming sessions as well as to serve as an indicator of the students' overall level of English proficiency. Comprehension test A (see Appendix I), comprehension test B (see Appendix J), and comprehension test C (see Appendix K), each featured 15 multiple-choice questions. In contrast to the vocabulary tests, the comprehension tests did not make use of any Japanese outside of the participation consent statement (see Appendix A), and the general test taking instructions of, "下の質問を答えてください。(Please answer the questions below)." In so doing, the students were required to understand the English used in the test as well as the in-game narrative content.

Just as with the vocabulary tests, all comprehension tests were conducted via Google Forms. The comprehension tests also made use of a multiple-choice format and required all participants to select a response before the test could be submitted. The test questions included matching the names of in-game characters to images (removed from the appendices), questions about in-game character relationships, and questions about in-game events. The multiple-choice selections were automatically randomized for each test generated. Just as with the vocabulary tests, all test items in the comprehension tests were derived from mandatory game sequences. The administration of the comprehension tests began from the spring semester of 2020 and continued to the end of the study. The only semester in which the comprehension tests were not administered was the fall semester of 2019. Since the comprehension test was not administered in the fall semester of 2019, the completion of all the comprehension tests was not required in order to be included for final analysis.

3.5.4 Survey structure

In addition to the comprehension tests, the LG group was also required to take pre-game and post-game surveys. Both pre- and post-game surveys aimed to collect both quantitative and qualitative data regarding the students' smartphone gaming habits, perceptions on the efficacy of smartphone gaming as a language-learning intervention, and perceptions of US culture. Both surveys were conducted entirely in Japanese and were, just as the vocabulary and comprehension tests, conducted via Google Forms. Both surveys utilized 10-point Likert scale questions, open-ended questions, and multiple-choice questions with optional respondent input. The pre-game survey (see Appendix E) measured: 1) how often students engage in smartphone gaming, 2) the language used when playing smartphone games, 3) where they play smartphone games, 4) experience with playing smartphone games in English, 5) how effectively they believe English smartphone gaming will improve English language proficiency, 6) what English language skills they believe will be improved by English smartphone gaming, 7) perceptions on US culture, 8) perceptions on the US educational system, 9) perceptions on students in the US, 10) what media sources they use to inform their opinions on US culture, 11) prior experience playing LIS and in what language, 12) their opinions on the advantages and disadvantages of utilizing smartphone gaming in formal EFL contexts, and 13) their opinions on how smartphone gaming in the EFL classroom compares to more traditional instructional methods. The post-game survey (see Appendix L) measured: 1) their willingness to play through the full game of LIS in English, 2) where they played LIS during the study program, 3) how well they comprehended the in-game narrative content, 4) what specifically was difficult to comprehend about said narrative content, 5) their willingness to play more smartphone games in the future, 6) what language they intend to use when playing smartphone games in the future, 7) how effectively they believe English smartphone gaming will improve English language proficiency, 8) how difficult the English used throughout the game was to comprehend, 9) what English language skills they believe will be improved by English smartphone gaming, 10) perceptions on US culture, 11) perceptions on the US educational system, 12) perceptions on students in the US, 13) what media sources they use to inform their opinions on US culture, 14) how much they perceived their own English proficiency to be improved by playing LIS in English, 15) their opinions on the advantages and disadvantages of utilizing smartphone gaming in formal EFL contexts, and 16) their opinions on how smartphone gaming in the EFL classroom compares to more traditional instructional methods.

3.6 Analysis procedures

The novel study design presented in this thesis makes use of both quantitative and qualitative analysis methods to interpret the data collected throughout the three-year study program. Due to the complex relationship between sociocultural factors and SLA (Lantolf, 1994), research on efficacious pedagogical methods within the field of SLA may benefit from mixed-method analysis (i.e., quantitative/qualitative analysis), identified as a desirable research method in social sciences to understand complex phenomena (Creswell, 1999). The present study therefore adopted a mixed-method data collection/analysis procedure to address the research questions.

Specifically, this study uses mixed-method analysis methods to shed light on how using a choose-your-own-adventure, single-player RPG such as LIS in formal EFL contexts may affect EFL vocabulary growth, perceptions on the efficacy of smartphone gaming for language learning, and perceptions of TL culture. Moreover, the mixed-method analysis methods used in this study also attempt to identify how specific within-group factors such as gender, gaming location, previous gaming experience, perceptions on the efficacy of smartphone gaming for language learning, prior LIS experience, level of comprehension of in-game narrative content/individual English language proficiency, individual gaming experience, and smartphone gaming habits may correlate to both vocabulary test scores and to one another. The quantitative analysis in this study also makes use of between-group analyses (i.e., between the LG and LC groups) in order to examine the efficacy of the LIS gaming intervention in conjunction with corrective feedback on EFL vocabulary acquisition and retention relative to the LC group which received only corrective feedback as an intervention.

3.6.1 Measuring EFL vocabulary acquisition and retention

In order to measure the EFL vocabulary acquisition and retention in the LG and LC groups, both within-group and between-group statistical analyses were performed on data gathered from the vocabulary tests. Specifically, averages from the vocabulary pre-test were compared to aggregate averages from vocabulary tests 1, 2, and 3 (immediate averages) before corrective feedback, delayed vocabulary post-tests (delayed post-test averages) after corrective feedback, and one-semester delayed vocabulary post-tests (OSD post-test averages) after corrective feedback. These statistical comparisons were used to establish whether there was an effect from the LIS gaming intervention relative to the LC group. An aggregate vocabulary post-test score of the immediate averages and the delayed post-test averages (total averages) from the LG group was also used to run statistical correlations to the comprehension test scores and items on the pre- and post-game surveys. The pre- and post-game surveys in the LG group were also used to identify factors (e.g., prior smartphone gaming experience in English) which could be used to help identify statistically significant predictive factors for vocabulary test scores. Finally, an individual question analysis (hereinafter referred to as IQA) was performed for both LG and LC groups to measure the percentage of correct answers for every test item from the vocabulary pre-test, immediate vocabulary tests, delayed vocabulary post-tests, and OSD post-tests. These IQA averages for each vocabulary test item were used to isolate the effects of each testing protocol on individual test items in order to analyze which test items are more or less receptive to each study group's testing protocol.

3.6.2 Measuring shifts in attitudes towards smartphone gaming

In order to measure the changes in perception regarding the efficacy of smartphone gaming as a language learning intervention, the data collected in the pre- and post-game surveys from the LG group were used to run statistical analyses. First, the averages of the 10-point Likert scale survey questions that featured in both pre- and post-game surveys were used to track perceptual changes in said pre and post paired questions. Second, for pre- and post-game multiple-choice survey questions probing which particular linguistic skills the participants believed smartphone gaming in English would improve the most, the number of responses for each multiple-choice selection were cataloged, whereby the changes in perception from pre- to post-game could be quantitatively tracked. Finally, the responses to open-ended questions that featured in both pre- and post-game surveys which probed for students' perceptions on the efficacy of smartphone gaming in formal EFL contexts were codified to assign numeric values to specific categories identified in the student responses, whereby the changes in perception could similarly be quantitatively tracked from pre- to post-game.

3.6.3 Measuring effects on the perception of US culture

In order to measure the changes in perception of US culture, the data from pre- and post-game surveys from the LG group were utilized to run statistical analyses similar to the methods described above. More specifically, the open-ended questions that featured in both pre- and post-game surveys which probed for students' perceptions on specific aspects of US culture that featured prominently in the narrative content of LIS (e.g., the US educational system) were codified to allow for quantitative tracking of statistically significant changes in perception from pre- to post-game. Utilizing the same method as described above, the students' responses were analyzed to isolate common categories. Responses with elements that fit under one or more of the identified categories imparted a score of one for each respective category. The final totals for each category were then compared between pre- and post-game surveys in order to ascertain how student perception had been/had not been affected by the narrative content of LIS. The following sections will detail the methods utilized to accomplish the statistical analyses stated above.

3.6.4 Measuring statistical significance

Within-group, paired, two-tailed *t*-tests were used in both LG and LC groups to establish if there were any statistically significant changes in individual test scores from pre-test averages to immediate, delayed post-test, and OSD post-test averages. Similarly, within-group, paired, two-tailed *t*-tests were used in the IQA for both LG and LC groups to establish if there were any statistically significant changes in percentages of correct answers for each vocabulary test item throughout the same testing schedule described above. Likewise, within-group paired, two-tailed *t*-tests were also used to establish if there were any statistically significant changes in the pre- and post-game survey measures described above. Between-group, two-sample, two-tailed *t*-tests were used to establish if there were statistically significant differences in the pre-test, immediate, delayed post-test, and OSD post-test averages between the LG and LC groups. Finally, between-group, two-sample, two-tailed *t*-tests and ANOVA analyses were used to compare the predictive factor groups probed in the pre- and post-game surveys from the LG group (e.g., prior smartphone gaming experience in English) to establish statistically significant differences in vocabulary test scores between each factor group within the LG group. The alpha level threshold for statistical significance used throughout the entire quantitative analysis was set at $p < 0.05$.

3.6.5 Measuring two-factor linear correlations

Correlation coefficients were used to establish linear relationships between vocabulary test scores, comprehension test scores, and pre- and post-game survey responses. The statistical significance of correlation coefficients were established by calculating the *t*-values and two-tailed *t*-distributions with two degrees of freedom. For correlation coefficients with alpha values of $p < 0.05$, linear regression analyses were performed to calculate the r^2 values in order to interpret the percentage of variance explained through linear correlations. For the present study, the strength of correlation for statistically significant *r*-values between 0.2 and -0.2 represented negligible statistical correlations, *r*-values > 0.2 and < -0.2 represented weak statistical correlations, *r*-values > 0.3 and < -0.3 represented moderate statistical correlations, *r*-values > 0.5 and < -0.5 represented moderately strong statistical correlations, and *r*-values > 0.7 and < -0.7 represented strong statistical correlations.

3.6.6 Measuring within-group variance

Analysis of variance (ANOVA) calculations were run to determine statistically significant, multi-factored variations within the LG group (e.g., smartphone gaming locations). Statistical significance was observed at $f > 3.11$.

3.6.7 Measuring skew and kurtosis

Skew and kurtosis were calculated to look for normal data distribution. After calculating the skew values, kurtosis values, and excess kurtosis for each of the vocabulary tests for both LG and LC groups (i.e., vocabulary pre-tests, immediate tests, delayed post-tests, and OSD post-tests) and for each of the comprehension tests for the LG group, *z*-tests were performed for both the skew and excess kurtosis coefficients. Normal distribution was observed for *z*-values < 3.29 and > -3.29.

3.6.8 Measuring effect size

The Cohen's *d* calculation was used to determine the effect size of the LG and LC group's respective testing protocols. The calculation was run on immediate tests, delayed post-tests, and OSD post-tests to examine how many standard deviations the means of the above-stated test averages were from the vocabulary pre-test means for the LG and LC groups. Using interpretations of Cohen's *d*-values from relevant literature to establish a baseline for the present study (Mcleod, 2019), Cohen's *d*-values between 0.2 and -0.2 represented negligible effects, *d*-values > 0.2 / < -0.2 represented a small effect size, *d*-values > 0.5 / < -0.5 represented a medium effect size, and *d*-values > 0.8 / < -0.8 represented a large effect size.

3.7 Summary of methodology

The methodology described in this chapter was used to address the four RQs: 1) How effective is the LIS gaming intervention at eliciting TL vocabulary acquisition and retention when used in conjunction with pre-determined testing protocols relative to a control group? 2) How do the gameplay mechanics in the smartphone application of LIS affect vocabulary acquisition and retention of TL vocabulary? 3) How will the LIS gaming intervention correlate to students' perceptions of and attitudes towards English smartphone gaming? and 4) How will the depictions of US culture in the narrative content of LIS affect how the students inform their impressions on US culture? The following chapters will present how the results of the final data analysis addresses each research question.

Chapter 4 Results for vocabulary acquisition and retention

The following chapter will detail the results of the data analysis which directly address RQ1: How effective is the LIS gaming intervention at eliciting TL vocabulary acquisition and retention when used in conjunction with pre-determined testing protocols relative to a control group? Specifically, this chapter will first present the results of the vocabulary test protocols from both LG and LC groups, analyze each group's vocabulary test data individually, then compare the results between the study groups. Finally, analyses of data distribution for each group's vocabulary test data will be used to determine how generalizable the vocabulary test results are to demographically similar populations.

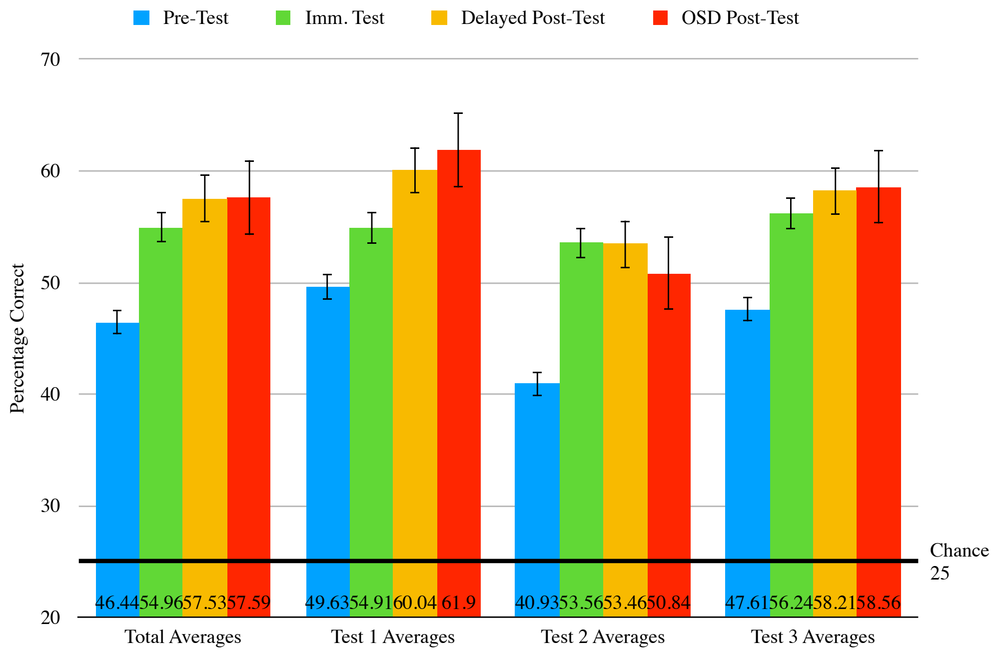
4.1 Modified vocabulary test results from the LG group

The modified vocabulary test results from the LG group (i.e., after removal of data non-compliant with the study program) are summarized in Table 4.1 and Figure 4.1. The testing procedure was summarized in Table 3.3. The following sections will highlight the most critical results from the gaming and vocabulary testing intervention in the LG group.

Table 4.1 Modified vocabulary test results from LG group

	Pre-test averages	Immediate test averages	Delayed post-test averages	OSD post-test averages
Sample size	*n*=78	*n* =78	*n* =78	*n* =23
Vocabulary test 1	49.63	54.91	60.04	61.9
Vocabulary test 2	40.93	53.56	53.46	50.84
Vocabulary test 3	47.61	56.24	58.21	58.56
Total averages	46.44	54.96	57.53	57.59
t-test for test 1		0.00087	0.00000011	0.0034
t -test for test 2		0.000000049	0.0000029	0.44
t -test for test 3		0.0000072	0.000021	0.0017
t-test for total avg.		0.000000000015	0.0000000069	0.0039
Cohen's *d* for test 1		0.38	0.64	0.74
Cohen's *d* for test 2		0.79	0.65	0.21
Cohen's *d* for test 3		0.64	0.67	0.86
C's *d* for total avg.		0.8	0.79	0.72

Figure 4.1 Modified vocabulary test results from LG group



On the whole, the gaming and vocabulary testing intervention in the LG group resulted in statistically significant improvements across all three testing protocols (i.e., immediate tests, delayed post-tests, and OSD post-tests) relative to pre-test averages. Concerning the immediate test averages, the 8.52 percentage point gain relative to pre-test averages suggests that the gameplay mechanics of LIS provided the students with the contextual information needed to answer more vocabulary test questions correctly without the aid of dictionaries or corrective feedback. The *p*-values of <0.001 calculated from paired, two-tailed *t*-tests also revealed that results from the immediate vocabulary tests were highly statistically significant. The strength of these results are further elaborated upon with a *d*-value of 0.8 for the immediate test averages indicating a strong effect of the LIS gaming intervention. Of particular note is how the difficulty of the individual pre-tests seemingly had comparatively little effect on the respective immediate test averages. The most difficult selection of vocabulary test items appeared in vocabulary test two, with a nearly seven percentage point lower average relative to test three and a nearly nine point lower average relative to test one in the pre-test values. However, this percentage gap narrows considerably in the immediate vocabulary tests. Another way of looking at the data would be to argue that the more difficult the vocabulary test items, the more the effective the gaming intervention was. This argument is also supported by the *d*-values for immediate tests one, two, and three, whereby the smallest gaming-intervention effect size of 0.38 (i.e., a relatively weak effect) was observed in immediate test one, the pre-test average of which was two points higher than pre-test three and nine points higher than pre-test two. In contrast to this relatively weak effect size indicated in the Cohen's *d* calculation, a moderate effect was observed in immediate test three (*d*-value of 0.64) and a borderline strong effect was observed in vocabulary test two (*d*-value of 0.79).

The gains in vocabulary scores observed in the immediate tests were further bolstered by the corrective feedback protocol in the testing intervention. Relative to pre-test averages, the *p*-value of <0.001 calculated with paired, two-tailed *t*-tests indicates that the delayed post-test averages are highly statistically significant. The Cohen's *d* value of 0.79 also indicates a borderline strong effect of the gaming and testing interventions. With a gain of 2.57 percentage points in the averages of the delayed vocabulary post-tests relative to the immediate tests, the corrective feedback testing intervention succeeded in eliciting only modest gains. However, these gains appear to remain stable across the vocabulary delay period and support the argument that the gaming and testing interventions in the LG group encourage long-term retention of vocabulary acquired within the study program. Indeed, the largest gains in the delayed post-tests were observed in vocabulary test one with a roughly five percentage point gain relative to the immediate test averages and a Cohen's *d* of 0.64 (a moderate effect) relative to the pre-test averages despite vocabulary test one having the longest delay (three weeks) in the delayed post-test category. In further support of the argument that the gaming and testing interventions in the LG group successfully elicit long-term retention of acquired vocabulary, further gains relative to the delayed post-tests were observed in the averages from the OSD post-tests with a Cohen's *d* of 0.72 (a moderate effect) relative to the pre-tests. With that being said, the differences in sample size (78 to 23) make any further comparisons challenging. The next section will therefore present the modified results from the 23-participant OSD dataset.

4.2 Modified OSD test results from the LG group

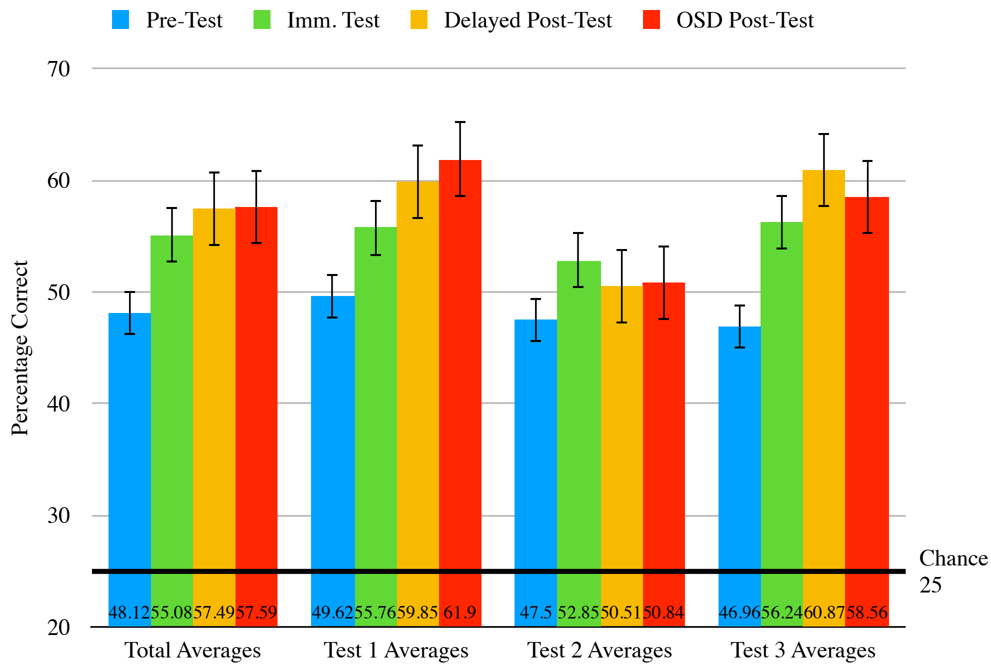
The modified vocabulary test results for the LG group participants (n=23) who took the OSD vocabulary post-test are summarized in Table 4.2 and Figure 4.2. Just as with the modified vocabulary test results from the LG group, paired, two-tailed t-tests were conducted for all three testing protocols to test for statistically significant gains relative to vocabulary pre-test values. Results showed that there were highly statistically significant gains in the vocabulary score averages across all three post-test protocols with p-values of <0.01 for all three testing protocols (i.e., immediate, delayed, and OSD). Furthermore, average Cohen's *d* values of 0.65 for the immediate tests and 0.72 for both delayed and OSD post-test groups indicate a moderately strong effect of the gaming intervention comparable to the full modified LG group sample. However, a notable exception to the statistically significant findings described above may be observed in the vocabulary test results from vocabulary test two across all four testing protocols.

Table 4.2 Modified OSD test results from LG group

	Pre-test averages	Immediate test averages	Delayed post-test averages	OSD post-test averages
Sample size	n=23	n =23	n =23	n =23
Vocabulary test 1	49.62	55.76	59.85	61.9
Vocabulary test 2	47.5	52.85	50.51	50.84
Vocabulary test 3	46.96	56.24	60.87	58.56
Total averages	48.12	55.08	57.49	57.59

t-test for test 1	0.0081	0.0015	0.0034
t -test for test 2	0.19	0.52	0.44
t -test for test 3	0.0014	0.00009	0.0017
t -test for total avg.	0.00028	0.001	0.0039
Cohen's *d* for test 1	0.41	0.63	0.74
Cohen's *d* for test 2	0.37	0.16	0.21
Cohen's *d* for test 3	0.69	1.06	0.86
C's *d* for total avg.	0.65	0.72	0.72

Figure 4.2 Modified OSD vocabulary test results from LG group



4.2.1 Anomalous results from OSD vocabulary test two items

In contrast to the results from the full modified LG group sample, the results from vocabulary test two in the modified OSD-only sample reveal non-significant gains in vocabulary score averages in both immediate and delayed post-test protocols. Similarly, the Cohen's *d* results from vocabulary test two in the OSD-only sample are considerably smaller in both immediate and delayed post-test protocols. Whereas the Cohen's *d* calculations revealed relatively strong effects of the gaming intervention for vocabulary test two in the full modified LG group sample, the Cohen's *d* values in the OSD-only sample revealed a weaker moderate effect in the immediate test (0.37) and a weak effect in the two-week delay post-test (0.16).

Of further interest is the drop in test score averages from 52.85 in immediate vocabulary test two to 50.51 in the two-week delay post-test. This drop occurred despite the corrective feedback protocol which provided participants with the answers to the tests after submitting the immediate tests. Although there is no definitive explanation as to why the results from vocabulary test two differ so significantly from the full modified LG group sample to the OSD-only sample, analyzing the data distribution of the vocabulary test two results may provide a plausible interpretation.

4.2.2 Data distribution from OSD vocabulary test two items

Given the relatively small sample size in the OSD-only sample (n=23) relative to the full modified LG group sample (n=78), the threshold for the number of participants to skew data is relatively low. When analyzing the data distribution of the vocabulary test two data from the OSD-only sample, the vocabulary pre-test data were found to be positively skewed (see Figure 4.3). The skew value of 0.78 and a z-score of 3.76 indicate statistically significant, moderate asymmetry in the OSD-only sample data for vocabulary pre-test two, meaning that the mode of the vocabulary pre-test two data from the OSD-only sample is lower than the median and mean of the dataset. When removing the five highest scores (LG23, LG28, LG61, LG92, and LG98) from the pre-test and immediate test two samples to restore symmetrical data distribution, results similar to the full modified LG group sample are observed with a pre-test two average of 43.16 and an immediate test two average of 51.28 (p=0.09, d=0.59, n=18). Another potential source of the anomalous results in the vocabulary test two data from the OSD-only sample relative to the full modified LG group sample is a highly statistically anomalous score of zero from participant LG10 in the two-week delay vocabulary post-test two. Although the kurtosis value of 2.2 and a z-score of -0.77 indicate that the outliers in the sample are within normal distribution (see Figure 4.4), when removing the extreme value from the two-week delay post-test dataset, the average jumps to 52.8, a much more comparable result to the full modified LG group sample results.

Figure 4.3 Modified OSD vocabulary test two data distribution (pre-test 2, imm. test 2)

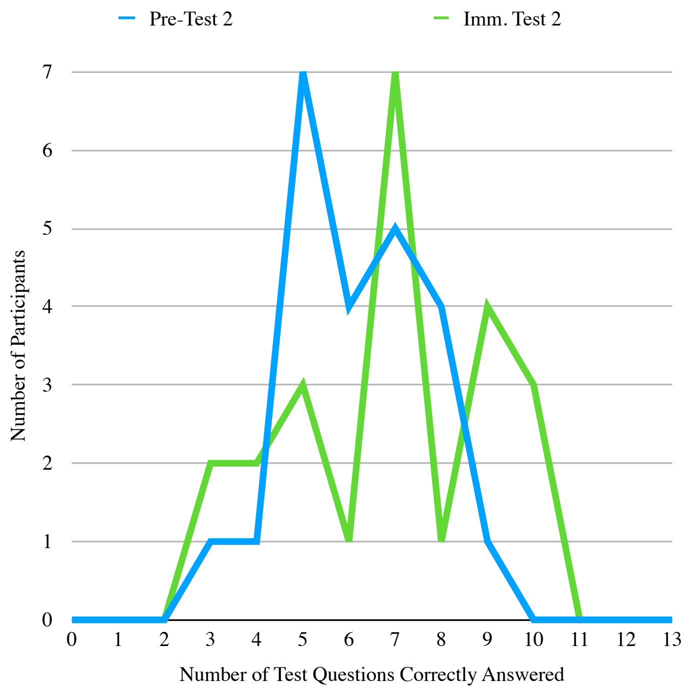
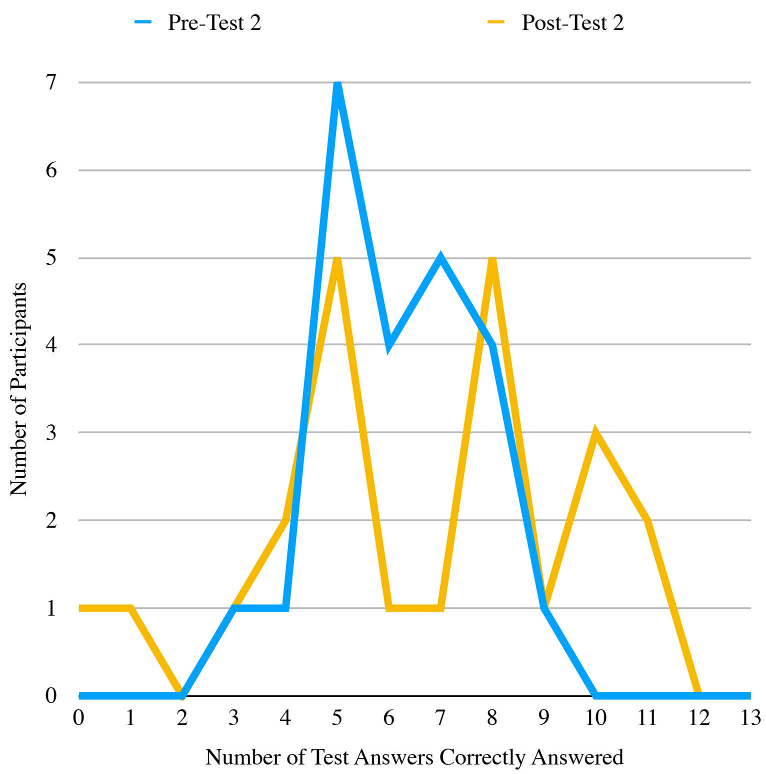
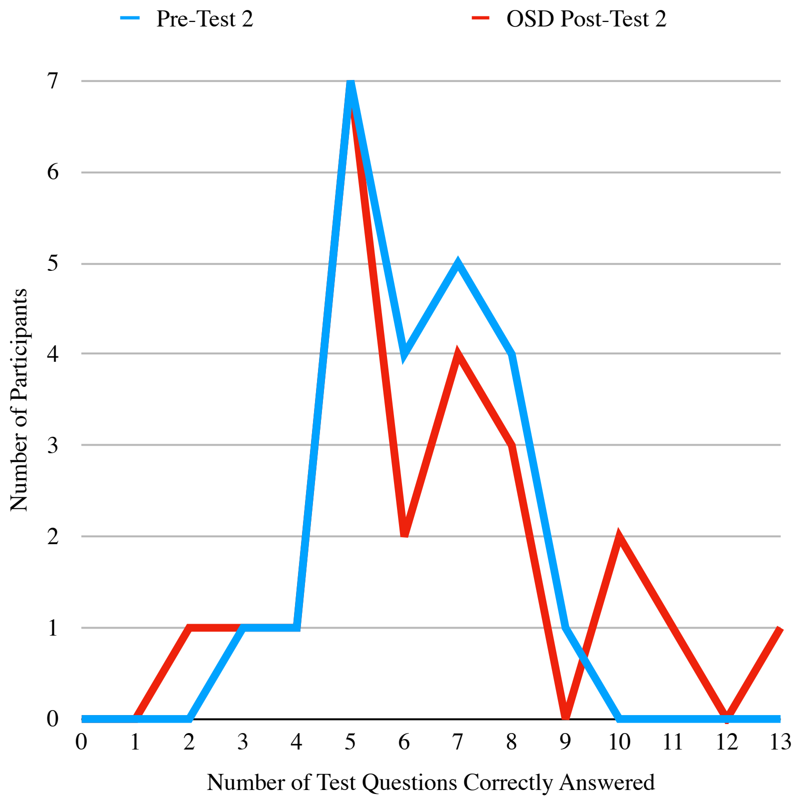


Figure 4.4 Modified OSD vocabulary test two data distribution (pre-test 2, post-test 2)



Although the arguable difficulty of vocabulary test two relative to the two other tests in the testing protocol may have played a role in the anomalous results from vocabulary test two in the OSD-only sample, other factors that may have contributed include: a mismatched sample in terms of overall English ability, a low sample size, and statistical anomalies in the test scores. However, even when taking these interpretations of the anomalous data from vocabulary test two in the OSD-only sample into consideration, the drop in scores in the OSD post-test two data is not explained by the interpretation above as removing the highest test scores did not significantly change the OSD values. On the whole though, the results in the OSD-sample are similar to the full modified LG group sample when totaling all three vocabulary tests in each testing protocol. The vocabulary test items in test two are arguably more difficult and less receptive to corrective feedback interventions, especially for long-term retention (see Figure 4.5). The argument that lower vocabulary test scores on the pre-test correlate to higher immediate test averages with the gaming intervention is nonetheless supported by the lower *p*-value and higher Cohen's *d* value that resulted after removing the highest scores in the dataset. As higher pre-test averages were observed to result in smaller Cohen's *d* values in the immediate test protocol, the data may suggest an upper threshold in the efficacy of the gaming intervention outlined in this study for demographically similar populations of participants. Overall, the significant results of the modified OSD-only group do add weight to the hypothesis that the gaming and testing interventions induced long-term retention of in-game vocabulary.

Figure 4.5 Modified OSD vocabulary test two data distribution (pre-test 2, OSD post-test 2)



4.3 Modified multiple play-through test results from the LG group

Comparisons of the modified vocabulary test results for the LG group participants who participated in the gaming intervention and testing protocols twice (*n*=8) are summarized in Table 4.3, Figure 4.6, Figure 4.7, Figure 4.8, and Figure 4.9. Specifically, the vocabulary test results from the first time through the gaming and testing interventions (hereinafter referred to as gaming/testing protocol one) are compared with the second time through (hereinafter referred to as gaming/testing protocol two). In order to analyze the efficacy of multiple play-throughs of the game and multiple exposures to the testing protocols, the vocabulary test averages were compared using paired, two-tailed *t*-tests to identify statistically significant shifts in test scores and Cohen's *d* calculations to identify any shifts in the effect size from gaming/testing protocol one to two. The paired, two-tailed *t*-test results and Cohen's *d* calculations for gaming/testing protocol one indicate the shifts in test scores from each testing protocol relative to the vocabulary pre-test values from gaming/testing protocol one. The paired, two-tailed *t*-test results and Cohen's *d* calculations for gaming/testing protocol two indicate the shifts in test scores from each testing protocol relative to their gaming/testing protocol one counterparts (i.e., vocabulary pre-test values from gaming/testing protocol two are compared directly with vocabulary pre-test values from gaming/testing protocol one and immediate vocabulary test values from gaming/testing protocol two are compared directly with immediate vocabulary test values from gaming/testing protocol one). The following sections will first detail the vocabulary test results from gaming/testing protocol one and then the comparison to gaming/testing protocol two.

Table 4.3 Modified multiple play-through test results from LG group

	Pre-test 1	Pre-test 2	Imm. test 1	Imm. test 2	D. post-test 1	D. post-test 2	OSD post-test 1	OSD post-test 2
Sample size	*n*=8	*n* =8	*n* =8	*n* =8	*n* =8	*n* =8	*n* =8	*n* =5
Vocab. test 1	41.92	58.09	47.06	60.3	51.48	76.48	58.09	65.89
Vocab. test 2	49.04	46.16	48.08	56.74	36.54	64.43	46.16	60
Vocab. test 3	41.67	52.5	48.34	60.84	51.67	68.34	52.5	62.67
Total averages	43.89	52.78	47.78	59.45	47.23	70.28	52.78	63.12
t-test for test 1		0.039	0.21	0.026	0.061	0.0017	0.039	0.82
t -test for test 2		0.53	0.89	0.25	0.075	0.03	0.53	0.21
t -test for test 3		0.024	0.23	0.035	0.0008	0.078	0.024	0.37
t -test for total avg.		0.037	0.017	0.011	0.26	0.0057	0.037	0.41
C's *d* for test 1		1.46	0.48	1.15	0.89	3.03	1.46	0.51
C's *d* for test 2		-0.24	-0.062	0.55	-0.79	1.42	-0.24	0.91
C's *d* for test 3		0.96	0.49	0.72	0.88	1	0.96	1.12
C's *d* total avg.		1.04	0.45	1.35	0.41	2.21	1.04	0.96

Figure 4.6 Modified multiple play-through vocabulary test results (pre-test values)

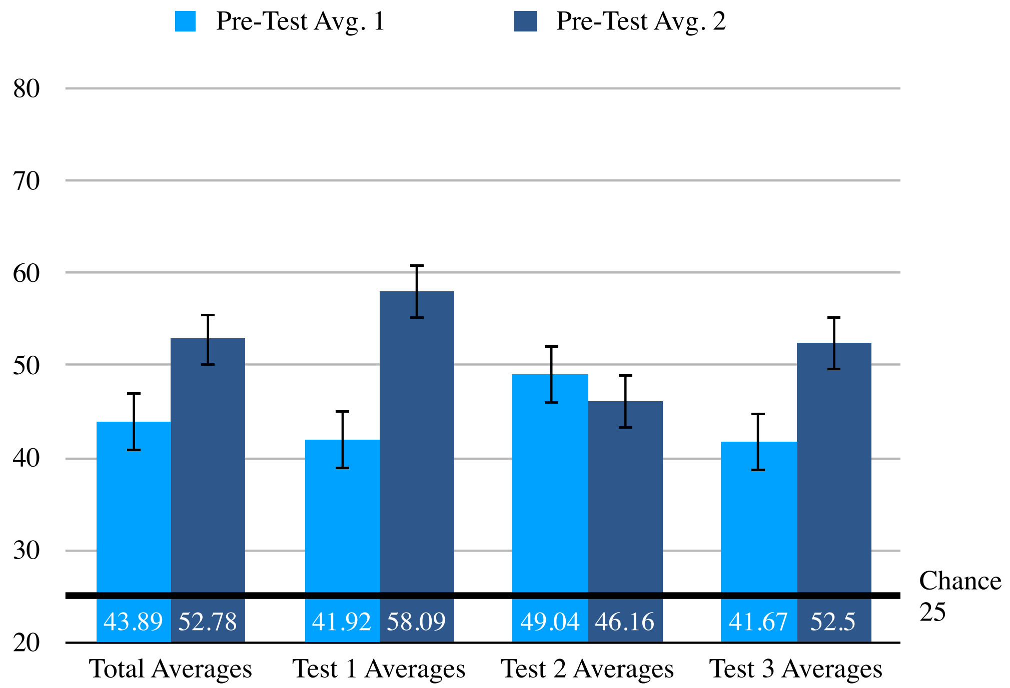


Figure 4.7 Modified multiple play-through vocabulary test results (immediate test values)

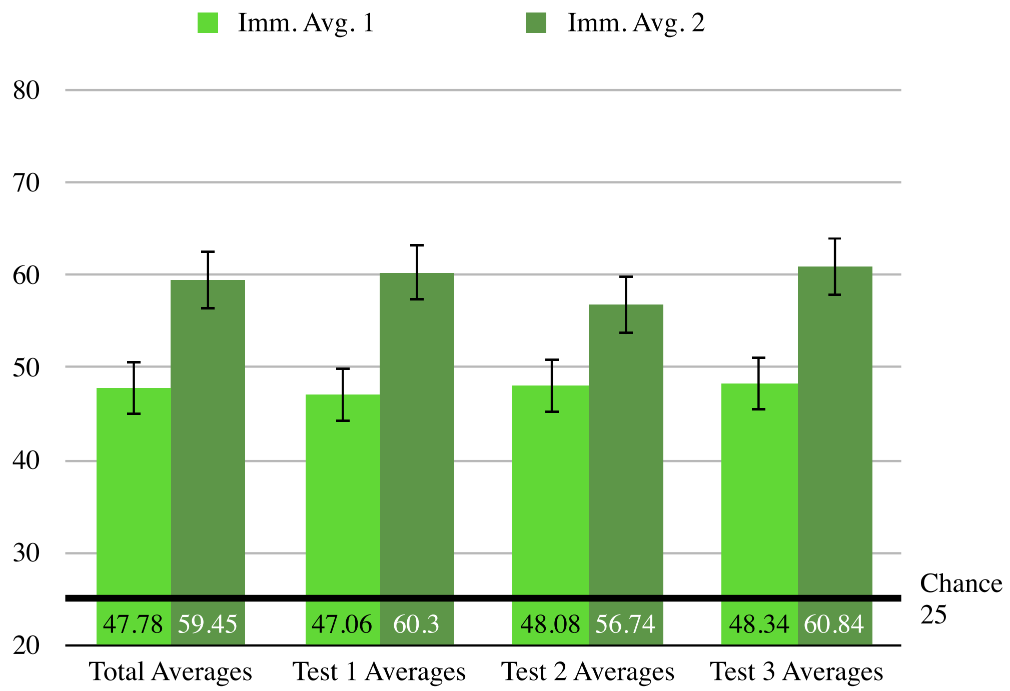


Figure 4.8 Modified multiple play-through vocabulary test results (delayed post-test values)

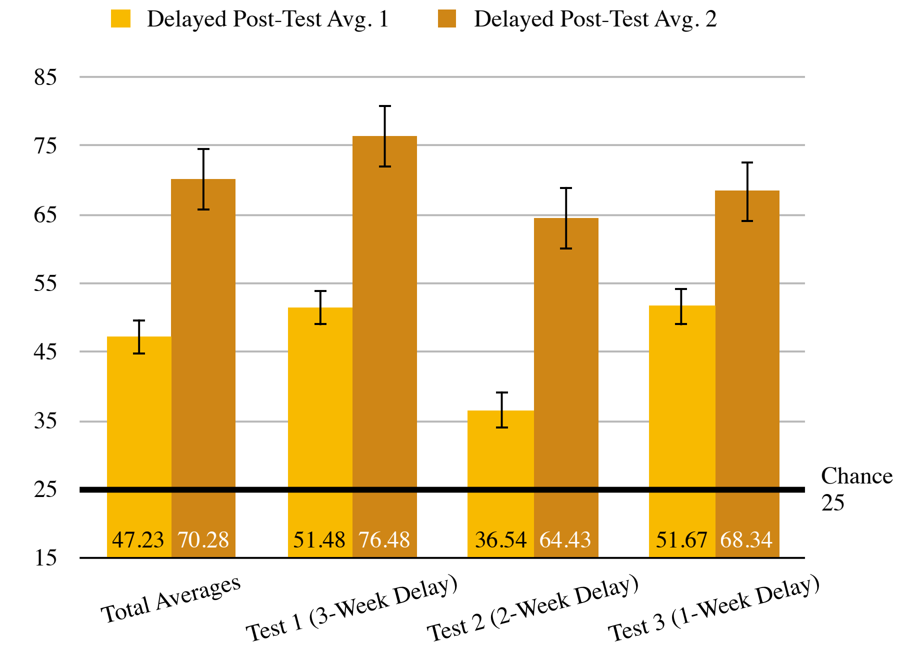
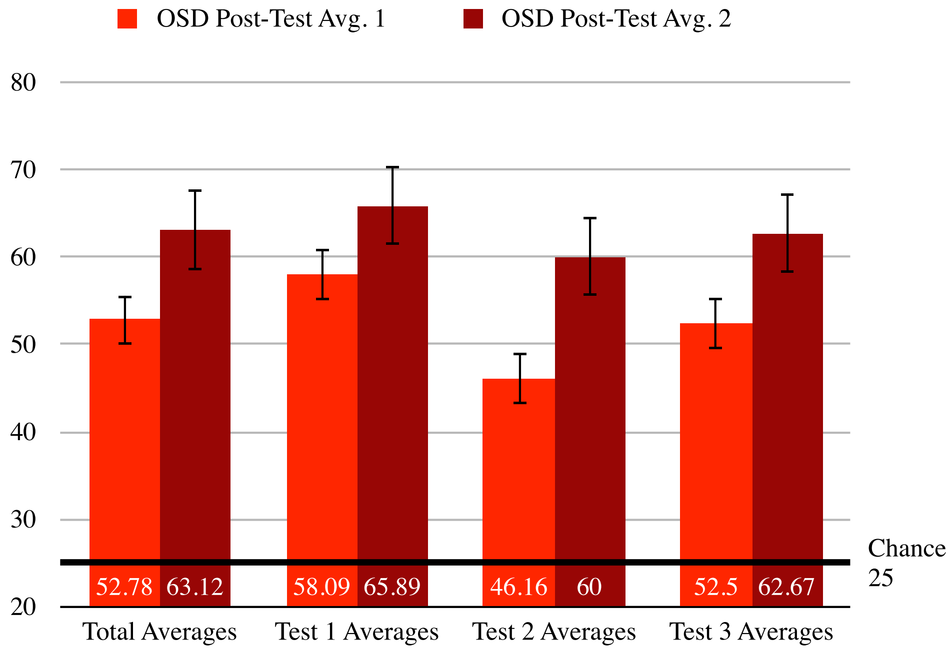


Figure 4.9 Modified play-through vocabulary test results (OSD post-test values)



4.3.1 Results from first play-through

The vocabulary test results from gaming/testing protocol one revealed statistically significant gains in total averages for both immediate and OSD post-test results relative to pre-test values. All gaming/testing protocol one *t*-test results for total averages were below the alpha threshold of <0.05 except for the total averages for the delayed post-test (p=0.26) and the Cohen's d values from the total averages in the immediate tests of 0.45 indicate moderate efficacy in the gaming and testing intervention for the multiple play-through sample (n=8) for the first play-through. The most striking gains in gaming/testing protocol one were actually observed in the OSD post-test total averages with a p-value of 0.037 and a Cohen's d value of 1.04 (strong effect). Although there were no statistically significant gains observed in the delayed post-test protocol, these resuts were skewed from the anomalous score of zero from participant LG10 as described in the previous section. With that being said, just as with the modified OSD-only sample analysis, vocabulary test two proved to be less receptive to the gaming and testing interventions in gaming/testing protocol one. Indeed, there were drops in scores relative to pre-test values in all three testing protocols for vocabulary test two, the drop in score in delayed post-test two itself approached statistical significance with a borderline strong negative effect size (p=0.075, d=-0.79). With the exception of vocabulary test two, the results were similar to the vocabulary test results from the full modified LG group sample. However, the small sample size (n=8) makes any comparisons inconclusive.

4.3.2 Results from second play-through

The results from the second play-through of the game and second round of tests from gaming/testing protocol two also demonstrated significant gains in vocabulary test averages, albeit with mixed statistical significance, and moderate to strong effect sizes. The only testing results in the second play-through to not achieve statistical significance were the total averages in the OSD post-test data from gaming/testing protocol two (n=5, p=0.41), albeit with a strong effect size (d=0.96). Of particular note are the vocabulary pre-test averages from gaming/testing protocol two as the 8.89 percent gain relative to the first play-through (p=0.037, d=1.04) is arguably the residual effect from gaming/testing protocol one. With that being said, the pre-test averages in vocabulary test two from gaming/testing protocol two were approximately three percentage points lower relative to those from gaming/testing protocol one. These results add further weight to the interpretation that the vocabulary test items in vocabulary test two are not only more difficult but are also less receptive to the gaming and testing interventions for long-term retention.

The immediate and delayed post-test results from gaming/testing protocol two revealed statistically significant results with strong effect sizes relative to their gaming/testing protocol one counterparts. The immediate test total averages from testing protocol two revealed an 11.67 point gain (p=0.011, d=1.35) relative to gaming/testing protocol one values, and the delayed post-test total averages from gaming/testing protocol two revealed a 23.05 point gain (p=0.0057, d=2.21) relative to gaming/testing protocol one values. The statistically significant, large percentage point gains with robust effect sizes in gaming/testing protocol two, especially in the immediate and delayed post-test protocols, provide convincing evidence that the efficacy of the gaming and testing interventions used in this study may be significantly increased with multiple exposures.

4.4 Modified vocabulary test results with multiple play-throughs removed from the LG group

Given that the results from the multiple play-through sample are significantly higher in gaming/testing protocol two relative to gaming/testing protocol one as defined in the section above, the possibility for the second play-through results from the eight participants who participated in the present study twice to positively skew the data from the full modified LG group sample will be addressed in this section. In order to ascertain whether the vocabulary test results from gaming/testing protocol two in the multiple play-through sample skewed the results from the full modified LG group sample, an analysis was conducted by removing the eight multiple play-through participants/one participant whose second play-through data remained after the removal of the first play-through data (hereinafter referred to as modified results two, n=69) and comparing these data with those from the full modified LG group sample (hereinafter referred to as modified results one, n=78). The results from this comparative analysis are summarized in Table 4.4, Figure 4.10, Figure 4.11, Figure 4.12, and Figure 4.13. In order to identify statistically significant differences between the two modified samples, two-sample, two-tailed t-tests were performed for each testing protocol (i.e., pre-test values from modified results one were compared to pre-test values from modified results two and immediate test values one to immediate test values two).

Table 4.4 Modified multiple play-through test results removed from LG group

	Pre-test 1	Pre-test 2	Imm. test 1	Imm. test 2	D. post-test 1	D. post-test 2	OSD post-test 1	OSD post-test 2
Sample size	n=78	n =69	n =78	n =69	n =78	n =69	n =23	n =18
Vocab. test 1	49.63	48.34	54.91	54.05	60.04	57.81	61.9	60.79
Vocab. test 2	40.93	40.03	53.56	53.07	53.46	52.07	50.84	48.3
Vocab. test 3	47.61	46.96	56.24	55.46	58.21	56.72	58.56	57.41
Total averages	46.44	45.48	54.96	54.24	57.53	55.79	57.59	56.05
t-test for test 1		0.55		0.73		0.47		0.85
t -test for test 2		0.7		0.87		0.72		0.69
t -test for test 3		0.75		0.74		0.62		0.81
t -test total avg.		0.53		0.71		0.55		0.77

Figure 4.10 Modified multiple play-through vocabulary test results removed (pre-test values)

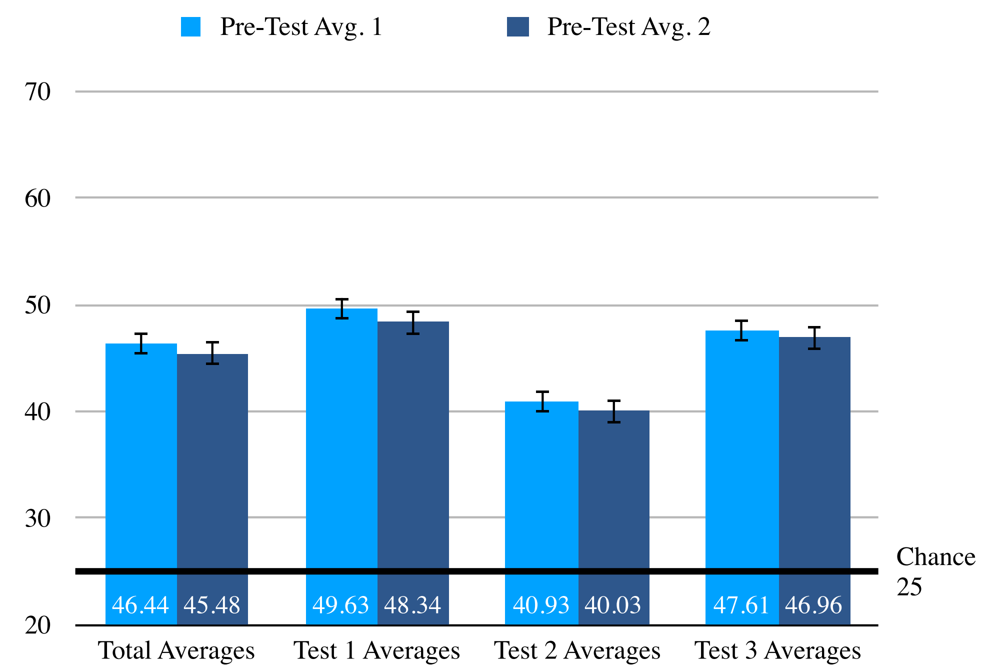


Figure 4.11 Modified multiple play-through vocabulary test results removed (immediate test values)

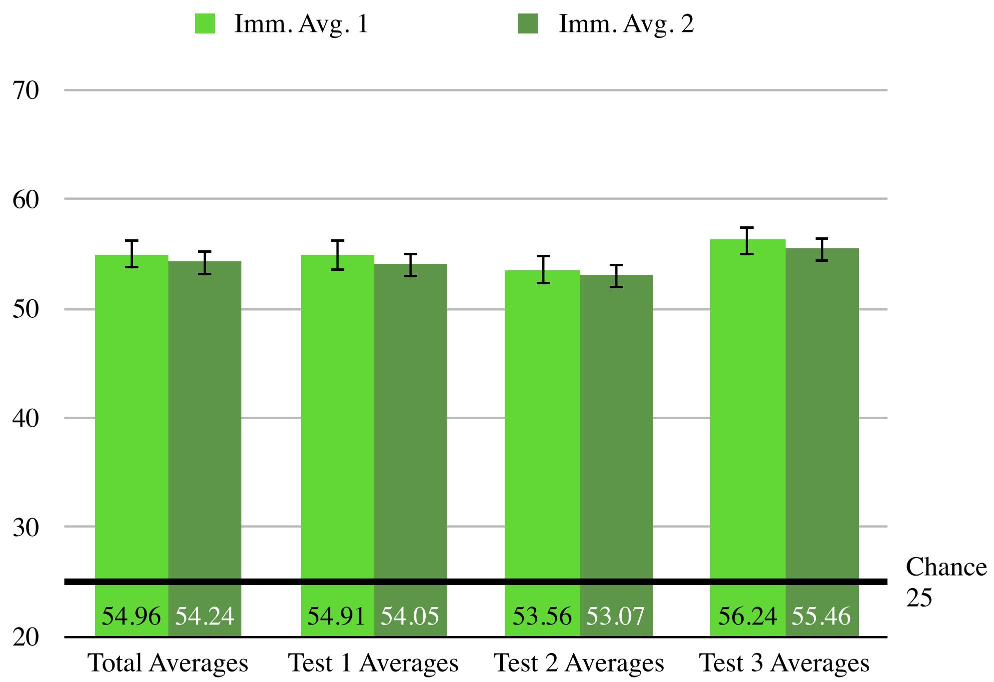


Figure 4.12 Modified multiple play-through vocabulary test results removed (delayed post-test values)

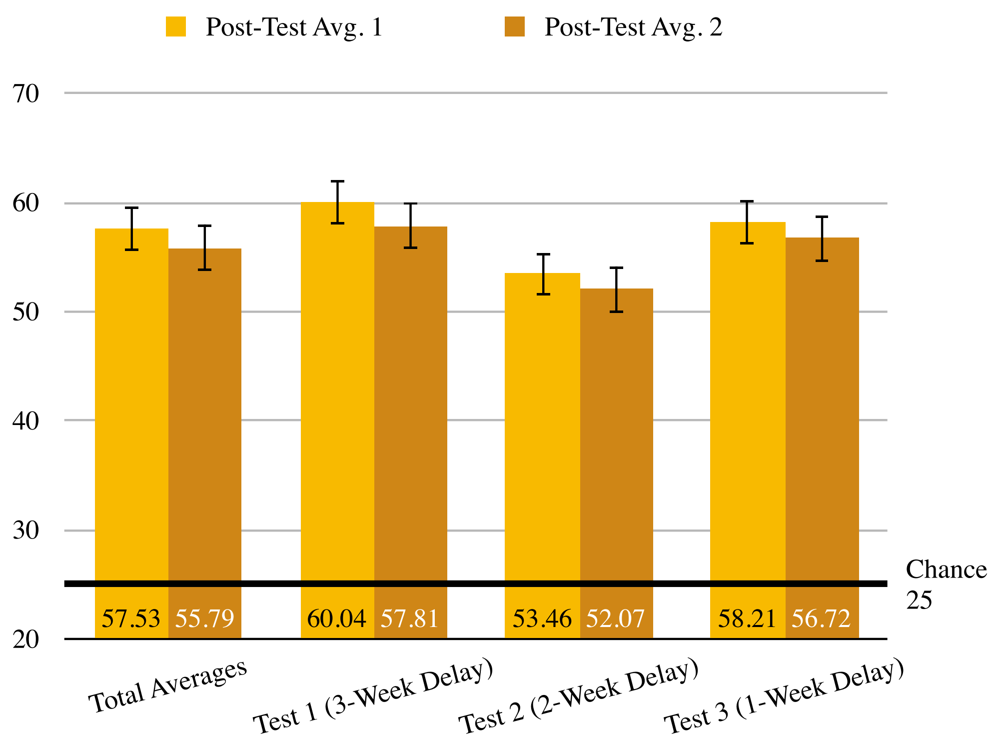
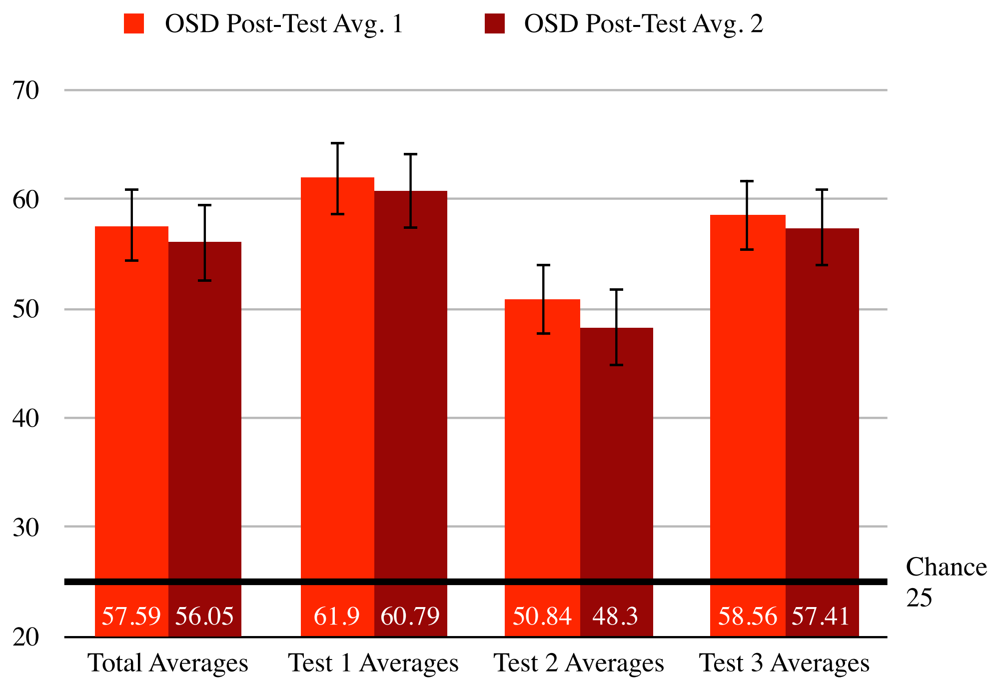


Figure 4.13 Modified multiple play-through vocabulary test results removed (OSD post-test values)



After removing the second play-through data from the eight multiple play-through participants in modified results one, the vocabulary test score averages in modified results two dropped slightly in each of the testing protocols. With that being said, none of the drops in score were statistically significant. Furthermore, even when analyzing the data from the individual vocabulary tests, none of the three vocabulary tests showed statistically significant changes when removing the second play-through data from the multiple play-through sample. Due to statistically negligible differences in total test score averages, the second play-through data from the multiple play-through sample were left in the full modified LG group sample for final analysis.

4.5 Analysis of vocabulary data distribution from the LG group

The distribution of data from the modified LG group is summarized in Table 4.5. The vocabulary test results from the modified LG group sample plotted with normal data distribution are summarized in Figure 4.14, Figure 4.15, Figure 4.16, and Figure 4.17. Each of the testing protocols were determined to conform to normal data distributions except for the delayed post-test data (see Figure 4.18), pre-test/delayed post-test three data (see Figure 4.19), and vocabulary pre-test one data (see Figure 4.20) which were determined to be approximately symmetrical. The non-normal, approximately symmetrical data distributions are analyzed in detail in the following section.

Table 4.5 Analysis of data distribution from LG group

	Pre-test values	Imm. test values	Post-test values	OSD post-test values
Sample size	*n*=78	*n*=78	*n*=78	*n*=23
Standard deviation test 1	13.12	14.8	18.73	18.68
Standard deviation test 2	14	17.77	23.55	20.05
Standard deviation test 3	12.43	14.56	18.51	14.49
St. dev. total averages	9.34	11.78	17.63	15.94
z-value for test 1 skew	5.68	-2.34	0.57	-2.23
z -value for test 2 skew	1.56	-0.15	-3.03	2.96
z -value for test 3 skew	-4.07	-2.28	-4.04	-0.48
z -val. for total skew avg.	1.89	1.22	-3.47	-1.4
z -value for test 1 kurtosis	-1.74	0.067	-1.17	-1.26
z -value for test 2 kurtosis	-0.58	-1.36	-1.32	-0.091
z -value for test 3 kurtosis	-0.44	-0.83	-0.69	1.63
z -val. for total kurt. avg.	-0.39	0.17	-1.09	-0.05

Figure 4.14 Pre-test normal data distribution from LG group

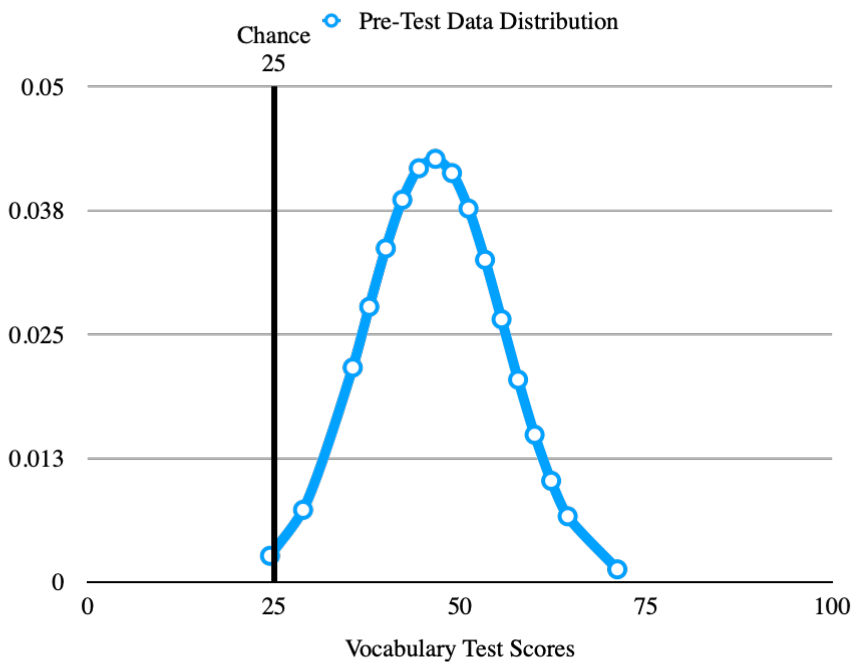


Figure 4.15 Immediate test normal data distribution from LG group

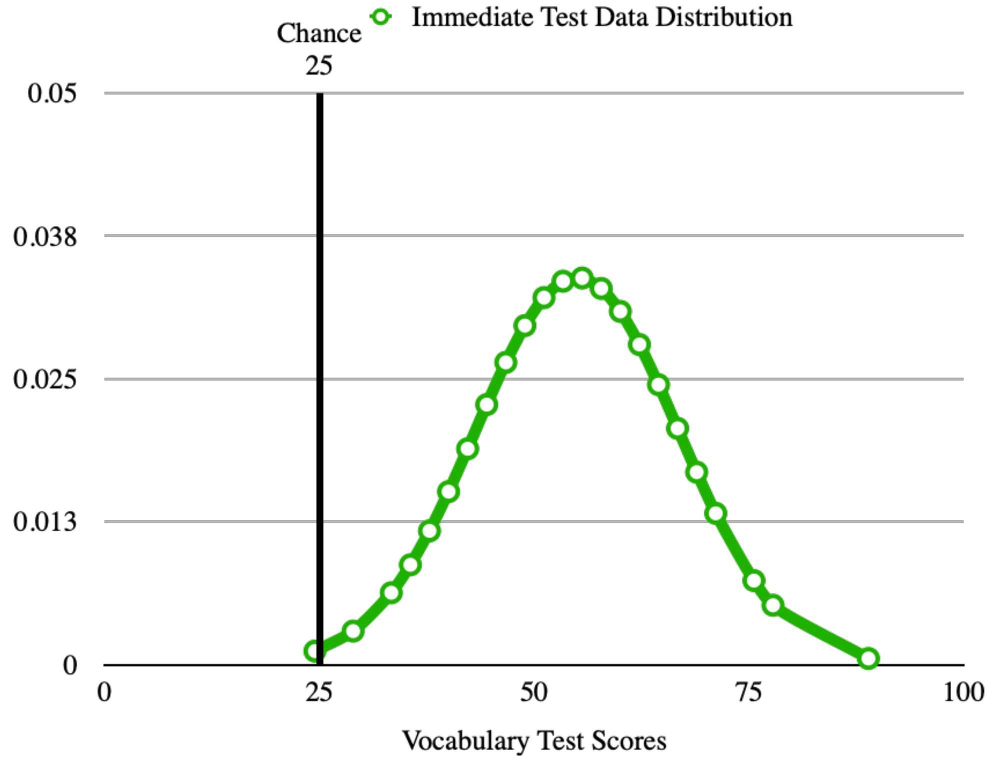


Figure 4.16 Delayed post-test normal data distribution from LG group

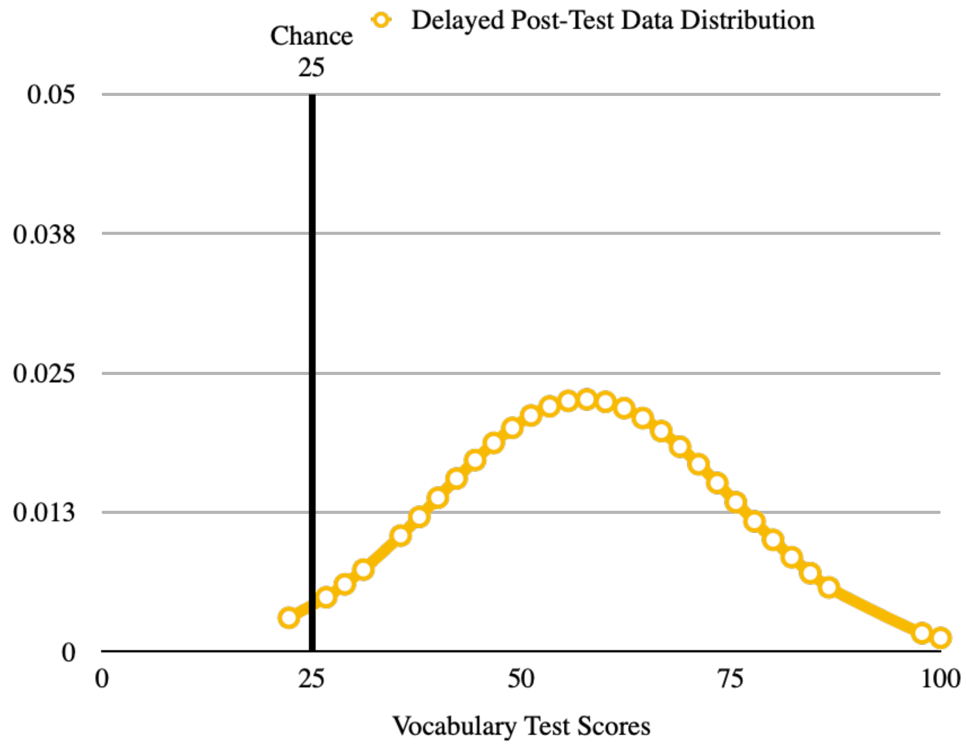


Figure 4.17 OSD post-test normal data distribution from LG group

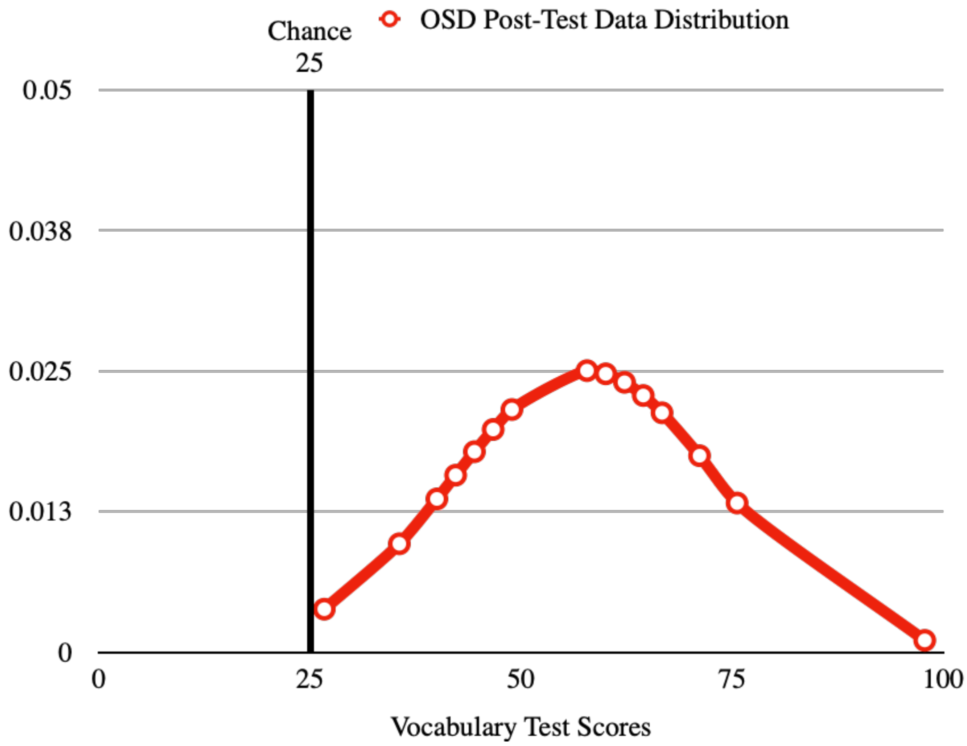


Figure 4.18 Comparative data distribution from LG group (pre-test and delayed post-test values)

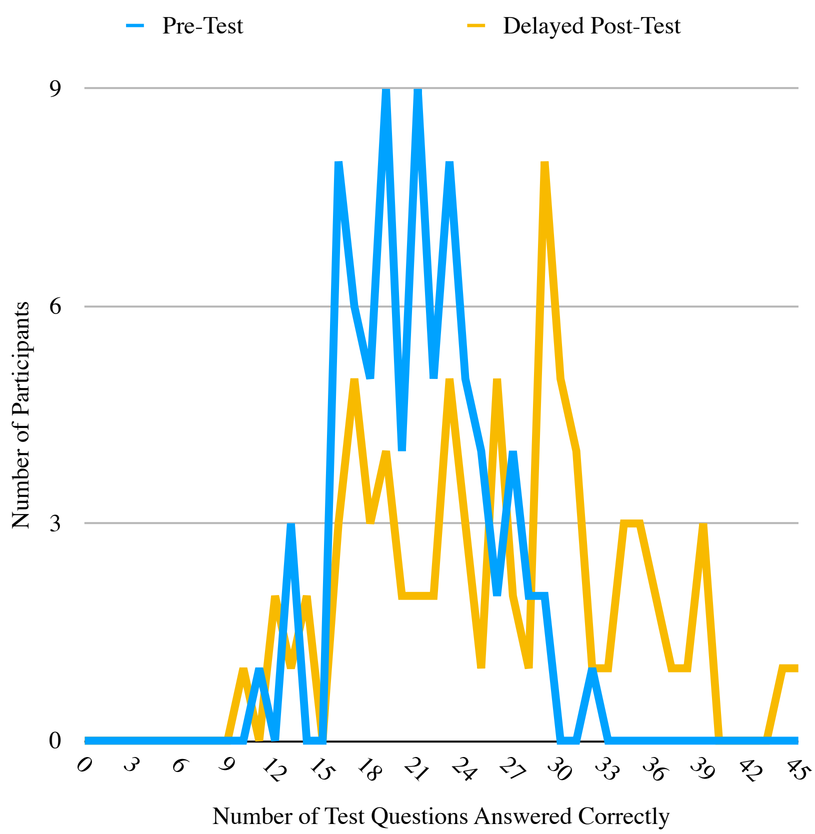


Figure 4.19 Comparative data distribution from LG group (pre-test 3 and delayed post-test 3 values)

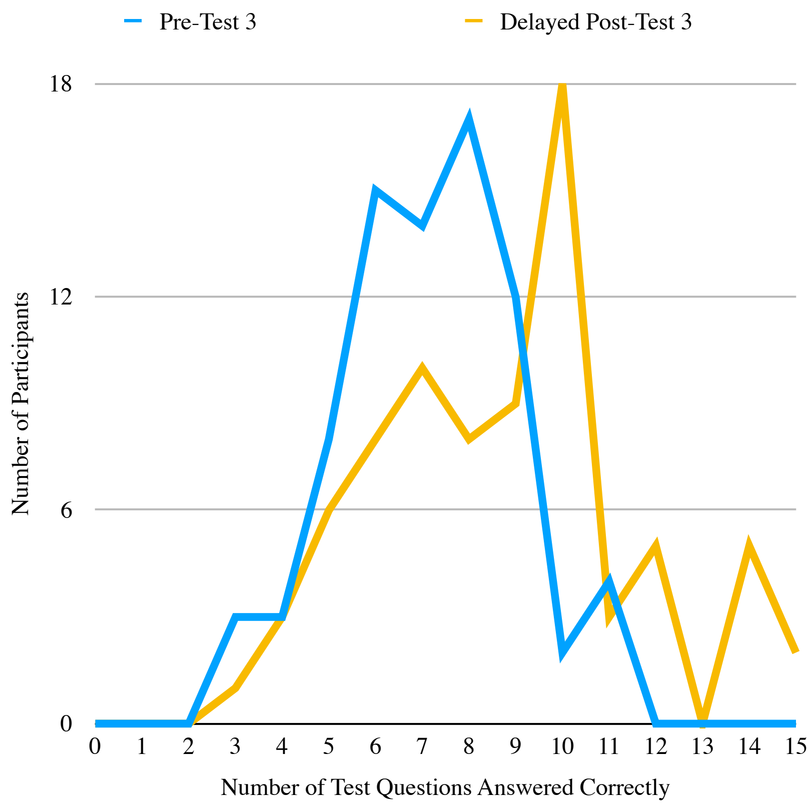
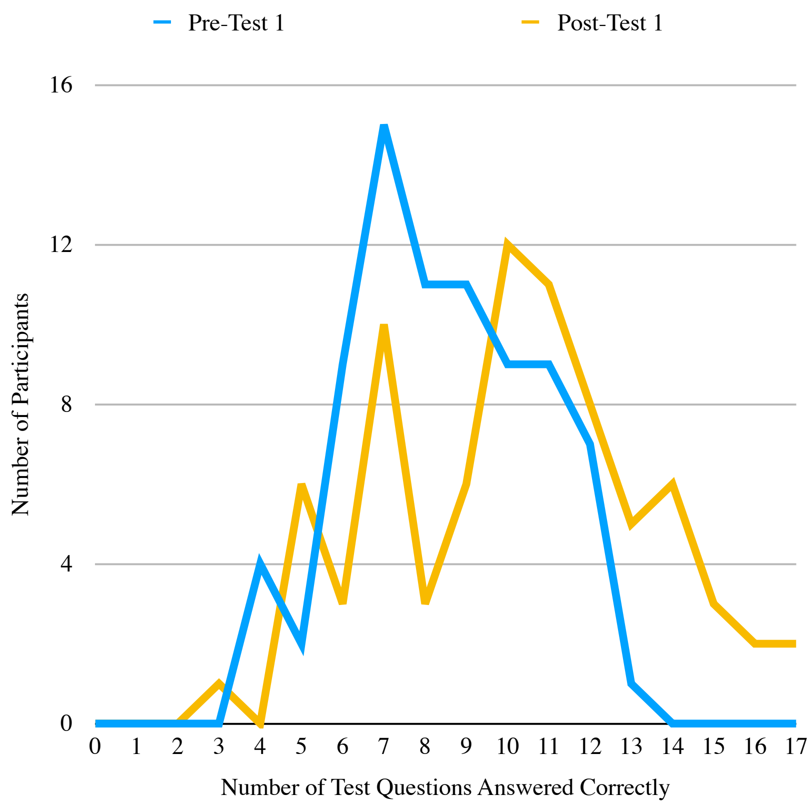


Figure 4.20 Comparative data distribution from LG group (pre-test 1 and delayed post-test 1 values)



4.5.1 Analysis of non-normal data distributions in the modified LG group sample

The data from the total averages in the delayed post-test protocol (see Figure 4.18) displayed a skew value of -0.39 and an excess kurtosis value of -0.67. The negative skew value is relatively low but nonetheless indicates slight asymmetry in data distribution, whereby a disproportionately large number of low test scores have pulled the median and mean down lower than the mode of the dataset. The excess kurtosis value is within normal data distribution parameters indicating a non-significant number of outliers. The primary source of this negative skew is coming from the data in delayed post-test three (see Figure 4.19). The skew value for delayed post-test three is -0.46 indicating a borderline moderate skew, whereby several low test scores have pulled the median and mean down lower than the mode of the dataset. The excess kurtosis value for delayed post-test three is -0.43 indicating a non-significant number of outliers in the dataset. The data from the delayed post-test protocol may therefore be described as being approximately symmetrical. However, the presence of negative skewness in some of the data also present the possibility that a number of participants in the modified LG sample group, albeit a statistically non-significant number, were nevertheless not as positively affected by the corrective feedback protocol as the majority in the same sample.

Another source of asymmetry in the data distribution from the modified LG group sample was observed in the data from vocabulary pre-test one (see Figure 4.20). In contrast to the negative skew seen in the delayed post-test protocol, the skew from vocabulary pre-test one displayed a positive skew of 0.64 indicating a moderate number of participants who scored significantly higher than the majority of participants in the sample, thus raising the median and mean of the vocabulary pre-test one sample above the mode of the dataset. The excess kurtosis value of -0.82 indicates however that there are not a significant number of outliers in the vocabulary pre-test one dataset. Due to the fact that this moderate asymmetry in the dataset is limited to vocabulary pre-test one and that the total averages from the pre-test protocol indicate approximate symmetry conforming to a normal data distribution, any subsequent rise in vocabulary pre-test averages caused by the positive skew is negligible. Overall, with the exceptions of low to moderate skews in data as described above, the data from the modified LG group sample conform well to normal data distribution models and provide limited evidence for the generalizability of the results of the gaming and testing interventions pursued in this study to demographically similar populations.

4.6 Analysis of OSD vocabulary data distribution from LG group

In addition to the data distribution analysis of the full modified LG group sample, the same analysis was conducted on the OSD-only sample from the LG group in order to ascertain the generalizability of the OSD-only data to demographically similar populations. These data are summarized in Table 4.6. For the purpose of determining the normality of the data distribution for each testing protocol, any *z*-value less/more than +/-1.96 was considered to be sufficient to accept the null hypothesis and assume an approximately symmetrical distribution of data, and hence normality. The *z*-value threshold to accept/reject the null hypothesis was set lower in the OSD-only sample due to the considerably smaller sample size (n=23). The vocabulary test results from the modified OSD-only sample plotted with normal data distribution are summarized in Figure 4.21, Figure 4.22, Figure 4.23, and Figure 4.24. Just as with the results from the full modified LG group sample, each of the testing protocols in the OSD-only sample were determined to conform to normal data distributions except for the delayed post-test total/three data (see Figure 4.25 and Figure 4.26), vocabulary pre-test/OSD post-test one data (see Figure 4.27), and vocabulary pre-test/OSD post-test two data (see Figure 4.5). As the data skew from vocabulary post-test two in the OSD-only sample was already analyzed in a previous section, the following section will focus on the remainder of the non-normal data distributions.

Table 4.6 Analysis of OSD data distribution from LG group

	Pre-test values	Imm. test values	Post-test values	OSD post-test values
Sample size	*n*=23	*n* =23	*n* =23	*n* =23
Standard deviation test 1	14.28	15.75	17.88	18.68
Standard deviation test 2	11.52	16.93	23.4	20.05
Standard deviation test 3	12.31	14.47	13.97	14.49
St. dev. total averages	9.49	11.8	15.8	15.94
z-value for test 1 skew	-3.09	-0.93	-0.25	-2.23
z-value for test 2 skew	3.76	-0.28	-0.25	2.96
z -value for test 3 skew	-1.19	-1.25	-1.99	-0.48
z -val. for total skew avg.	-0.39	-1.55	-2.11	-1.4
z-value for test 1 kurtosis	-2.27	-1.1	-0.45	-1.26
z -value for test 2 kurtosis	-0.86	-0.63	-0.77	-0.091
z -value for test 3 kurtosis	-0.74	-0.52	-0.23	1.63
z -val. for total kurt. avg.	-1.21	-1.59	-0.58	-0.05

Figure 4.21 Pre-test normal data distribution from OSD-only sample

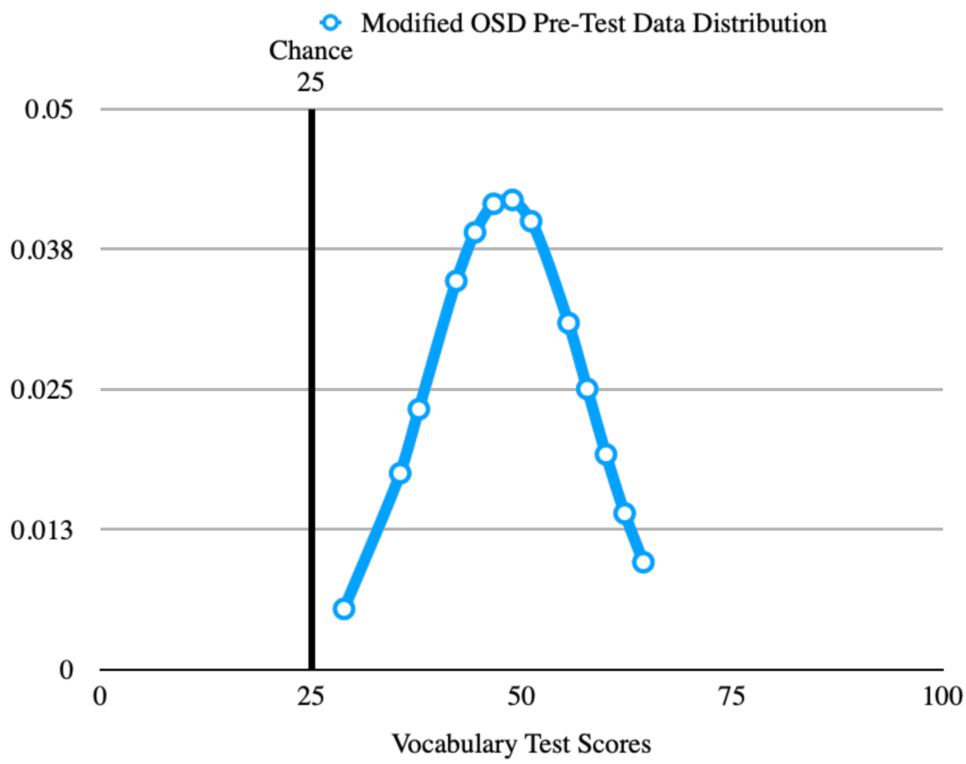


Figure 4.22 Immediate test normal data distribution from OSD-only sample

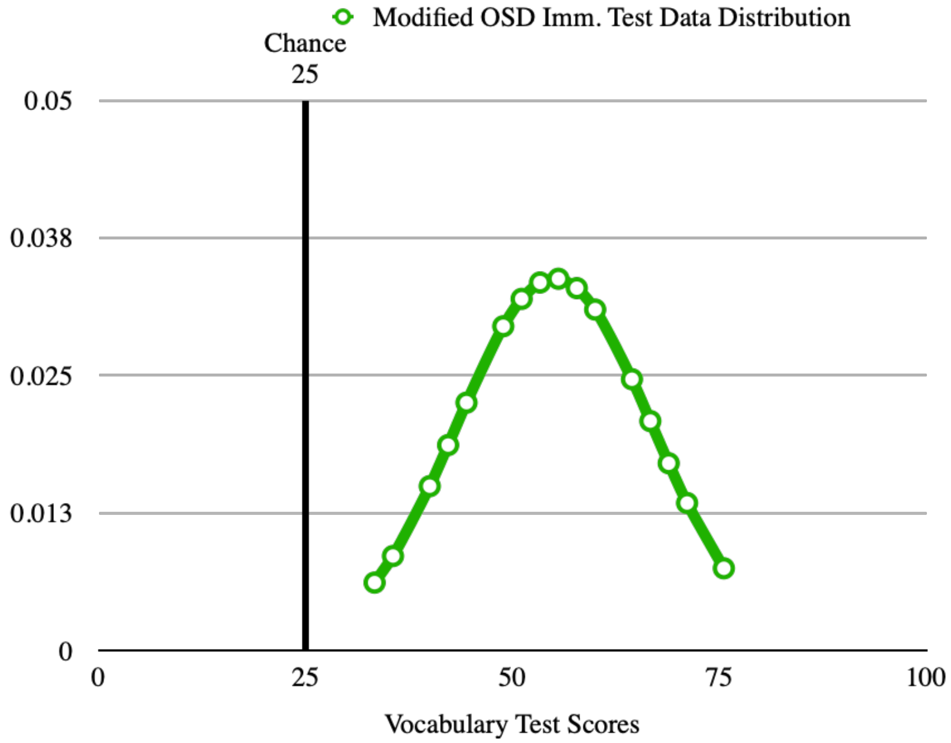


Figure 4.23 Delayed post-test normal data distribution from OSD-only sample

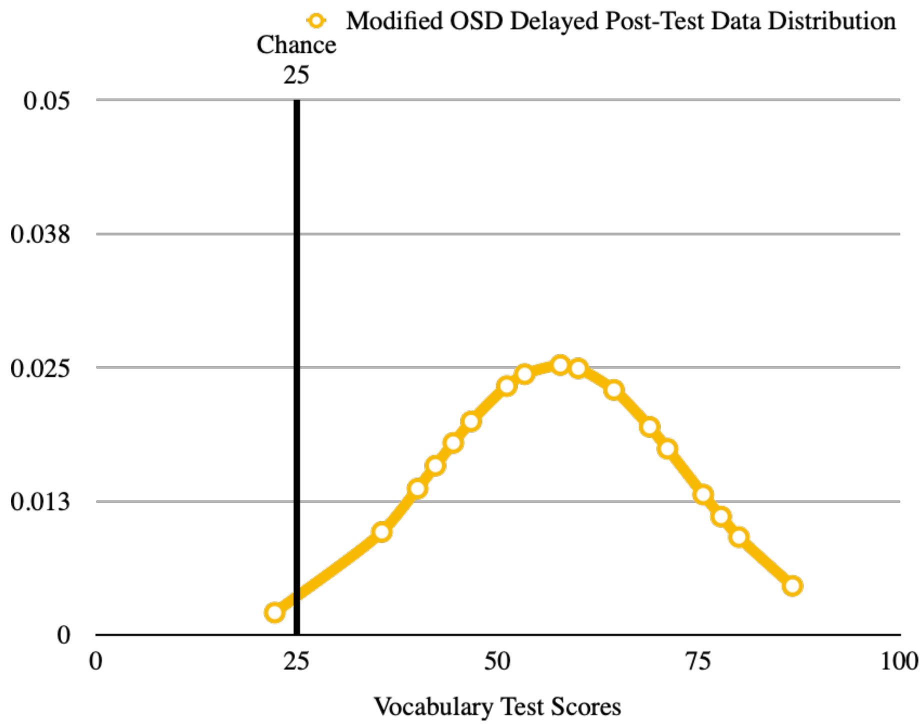


Figure 4.24 OSD post-test normal data distribution from OSD-only sample

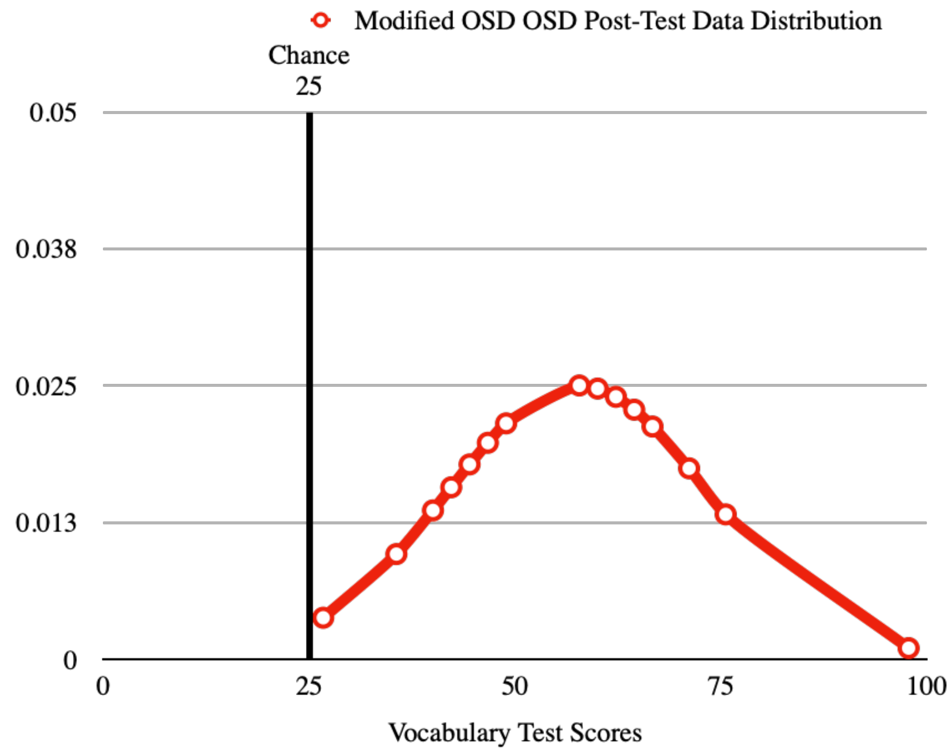


Figure 4.25 Comparative data distribution from OSD-only sample (pre-test and delayed post-test values)

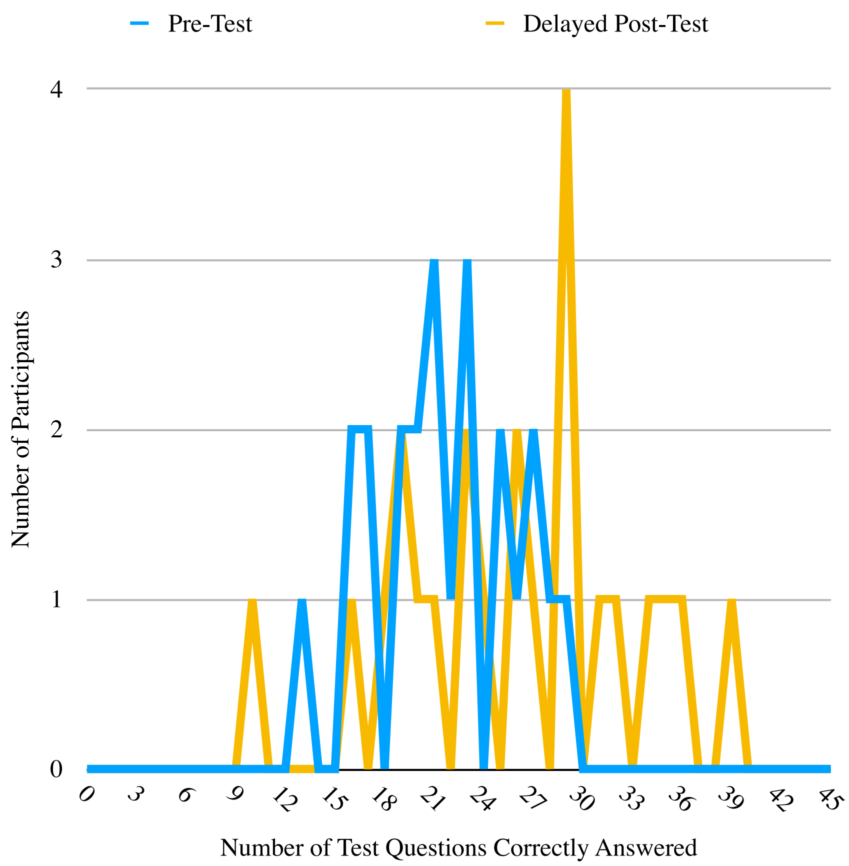


Figure 4.26 Comparative data distribution from OSD-only sample (pre-test 3 and delayed post-test 3 values)

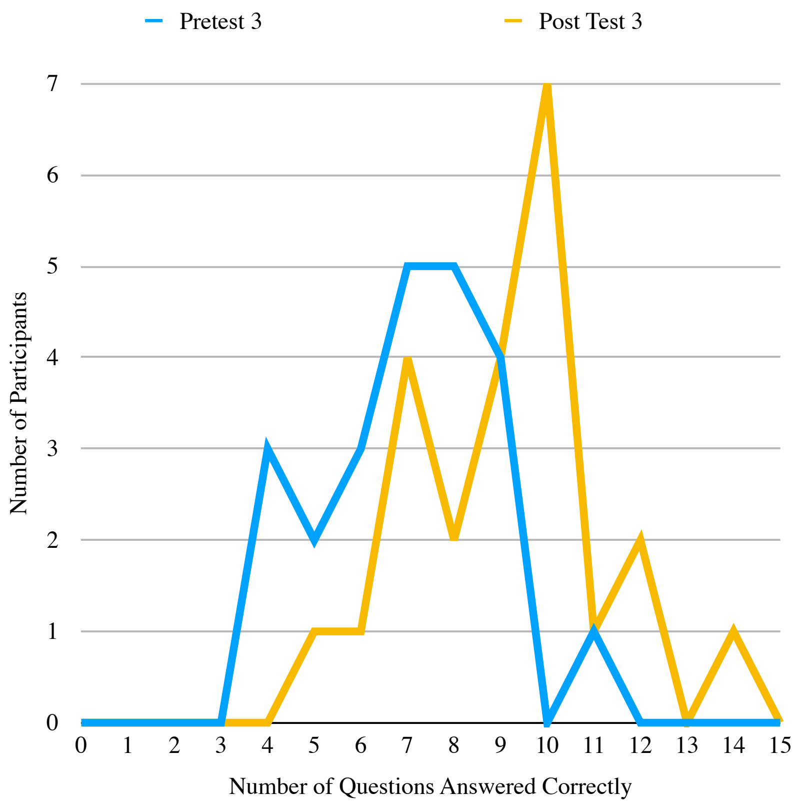
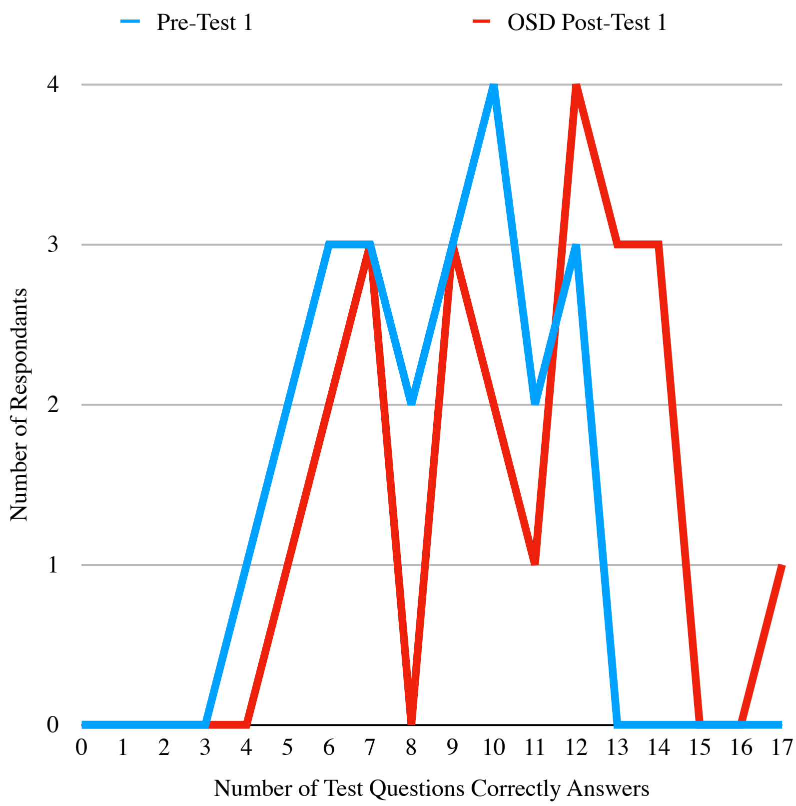


Figure 4.27 Comparative data distribution from OSD-only sample (pre-test 1 and OSD post-test 1 values)



4.6.1 Analysis of non-normal data distribution in OSD-only sample from LG group

Just as with the full modified LG group sample, the delayed vocabulary post-test protocol in the OSD-only sample displayed a negative skew (see Figure 4.25). With a skew value of -0.44, the data may be described as displaying a small number of anomalously low test scores which have dropped the median and mean below the mode of the dataset. The excess kurtosis value of -0.67 indicates that there are not a significant number of outliers in the dataset. Again, just as with the full modified LG group sample, the source of asymmetry in the dataset stems from vocabulary post-test three (see Figure 4.26). With a skew value of -0.42 and an excess kurtosis value of -0.32, the same interpretation of the data used with total averages for the delayed post-test protocol may be applied to the vocabulary post-test three dataset. Although the OSD-only sample is a modified group sample from the LG group and hence the data come from the same source, the significant non-symmetrical results in the same delayed vocabulary post-tests even in a smaller sample add weight to the interpretation that not all participants were responding to the corrective feedback protocol as effectively as the grand majority of other participants, especially with regard to delayed post-test three. With that being said, the level of skew observed in the delayed post-test protocol is relatively low, and the data distribution from this particular testing protocol still remains approximately symmetrical.

In contrast to the vocabulary pre-test one data from the full modified LG group sample which displayed a positive skew, the vocabulary pre-test one data from the OSD-only sample displayed a negative skew (see Figure 4.27). With a skew value of -0.65, the level of skew may be described as moderate, whereby a number of anomalously low test scores have dropped the median and mean below the mode of the dataset. Furthermore, the significantly low excess kurtosis value of -1.27 indicates a leptokurtic distribution (i.e., an absence of outliers). The exact source of this asymmetry is difficult to generalize, particularly considering that the same test in the same test protocol in the full modified LG group sample displayed a negative skew. Furthermore, a negative skew was also observed in the OSD post-test one dataset from the OSD-only sample (see Figure 4.27). With that being said, the skew value of -0.47 indicates a low level of skew, and the excess kurtosis value of -1.04 puts the data into the normal data distribution range for outliers. In a general sense, it may be concluded that a number of students in the OSD-only sample had more difficulty with the vocabulary test items in vocabulary test one relative to the majority of participants from the full modified LG group sample and that the gaming and testing interventions (including corrective feedback) helped to correct some of the data skew. With the exception of the negative skew in the total averages of the delayed post-test protocol, the OSD-only sample also displayed approximately symmetrical data distribution and adds weight to the interpretation that the results from the gaming and testing interventions used in the present study may apply to other demographically similar populations.

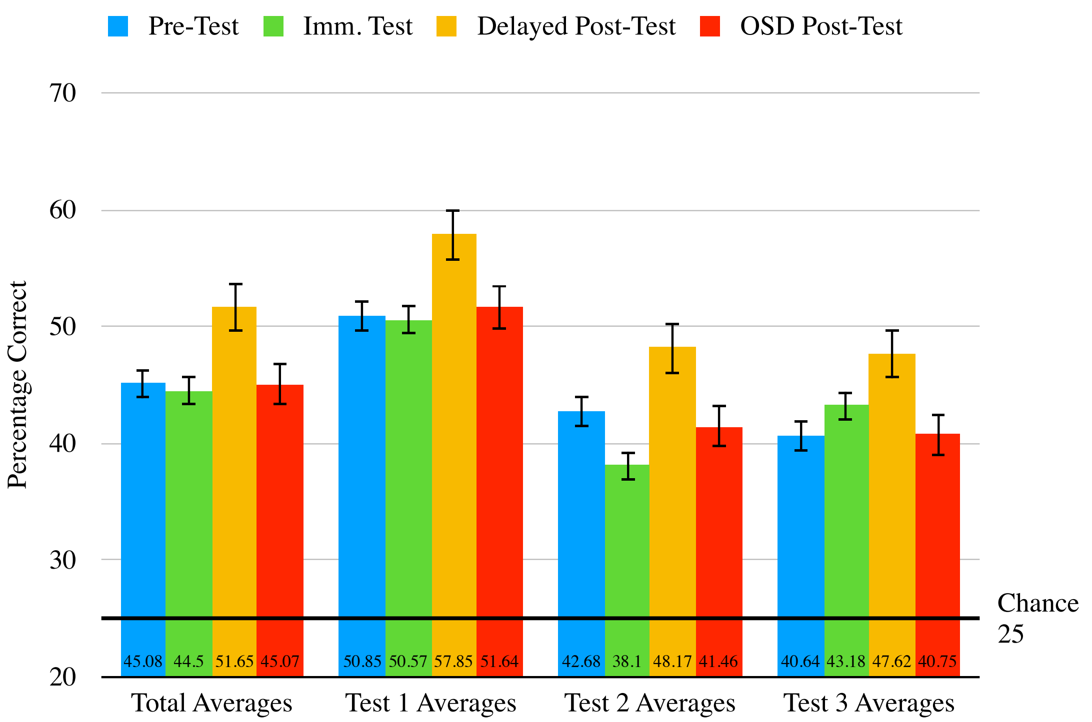
4.7 Modified vocabulary test results from LC group

The modified results from the LC group are summarized in Table 4.7 and Figure 4.28. Except for the gaming intervention, the same testing protocol and data analysis methods were used in the LC group. On the whole, there were only statistically significant gains in the delayed post-test protocol after the presentation of corrective feedback upon completion of the immediate tests. The p-value of 0.0026 and Cohen's d value of 0.57 indicate a weaker effect relative to the delayed post-test protocol in the LG group, but a statistically significant moderate effect of the LC group testing intervention, nonetheless. The immediate test and OSD post-test protocols resulted in a total drop of 0.58 percent and 0.01 percent respectively. The total drop in the OSD post-test protocol provides limited evidence for the lack of long-term retention of the gains observed in the delayed post-test protocol. In order to determine whether the vocabulary test results observed in the LC group are skewed, the following section will analyze the data distribution of each of the testing protocols.

Table 4.7 Modified vocabulary test results from LC group

	Pre-test averages	Immediate test averages	Delayed post-test averages	OSD post-test averages
Sample size	*n*=42	*n*=42	*n*=42	*n*=18
Vocabulary test 1	50.85	50.57	57.85	51.64
Vocabulary test 2	42.68	38.1	48.17	41.46
Vocabulary test 3	40.64	43.18	47.62	40.75
Total averages	45.08	44.5	51.65	45.07
t-test for test 1		0.92	0.023	0.69
t -test for test 2		0.065	0.097	0.93
t -test for test 3		0.35	0.033	0.16
t -test for total avg.		0.68	0.0026	0.5
Cohen's *d* for test 1		-0.02	0.52	0.058
Cohen's *d* for test 2		-0.33	0.32	-0.079
Cohen's *d* for test 3		0.19	0.4	0.0085
C's *d* for total avg.		-0.069	0.57	-0.0017

Figure 4.28 Modified vocabulary test results from LC group



4.8 Analysis of vocabulary data distribution from LC group

In order to ascertain comparability for the data collected from the modified LC group sample to the modified LG group sample, the normality of data distribution was analyzed by calculating skew, kurtosis, and the subsequent *z*-values for said statistics. These data are summarized in Table 4.8. For the purpose of determining the normality of the data distribution for each testing protocol, any *z*-value less/more than +/-1.96 was considered to be sufficient to accept the null hypothesis and assume an approximately symmetrical distribution of data. The vocabulary test results from the modified LC group sample plotted with normal data distribution are summarized in Figure 4.29, Figure 4.30, Figure 4.31, and Figure 4.32. Each of the testing protocols were determined to conform to normal data distributions except for the delayed post-test data (see Figure 4.33, Figure 4.34, and Figure 4.35), vocabulary pre-test/immediate test three data (see Figure 4.36), and OSD post-tests two and three (see Figure 4.37 and Figure 4.38). The non-normal data distributions are analyzed in detail in the following section.

Table 4.8 Analysis of data distribution from LC group

	Pre-test values	Imm. test values	Post-test values	OSD post-test values
Sample size	*n*=42	*n* =42	*n* =42	*n* =18
Standard deviation test 1	12.11	16.02	14.58	15.31
Standard deviation test 2	15.77	12.01	18.22	15.19
Standard deviation test 3	13.48	13.19	20.49	11.41
St. dev. total averages	8.59	8.16	13.96	11.78
z-value for test 1 skew	-1.13	-0.96	-3.05	1.27
z -value for test 2 skew	1.73	-0.2	-2.02	2.98
z -value for test 3 skew	3.51	-4.99	2.41	2.76
z-val. for total skew avg.	-1.2	-0.25	4.03	1.02
z-value for test 1 kurtosis	0.41	0.21	0.11	0.3
z-value for test 2 kurtosis	0.093	-0.52	0.036	-0.19
z-value for test 3 kurtosis	-1.84	-1.26	-0.37	-0.52
z-val. for total kurt. avg.	-0.65	1.43	-0.21	0.044

Figure 4.29 Pre-test normal data distribution from LC group

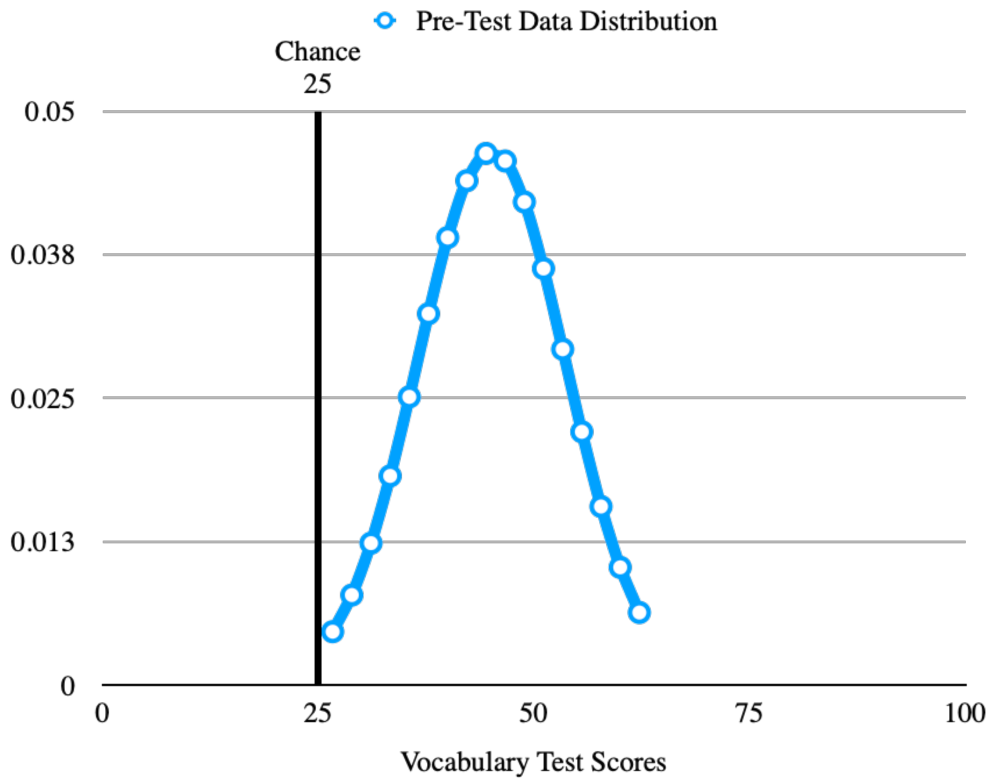


Figure 4.30 Immediate test normal data distribution from LC group

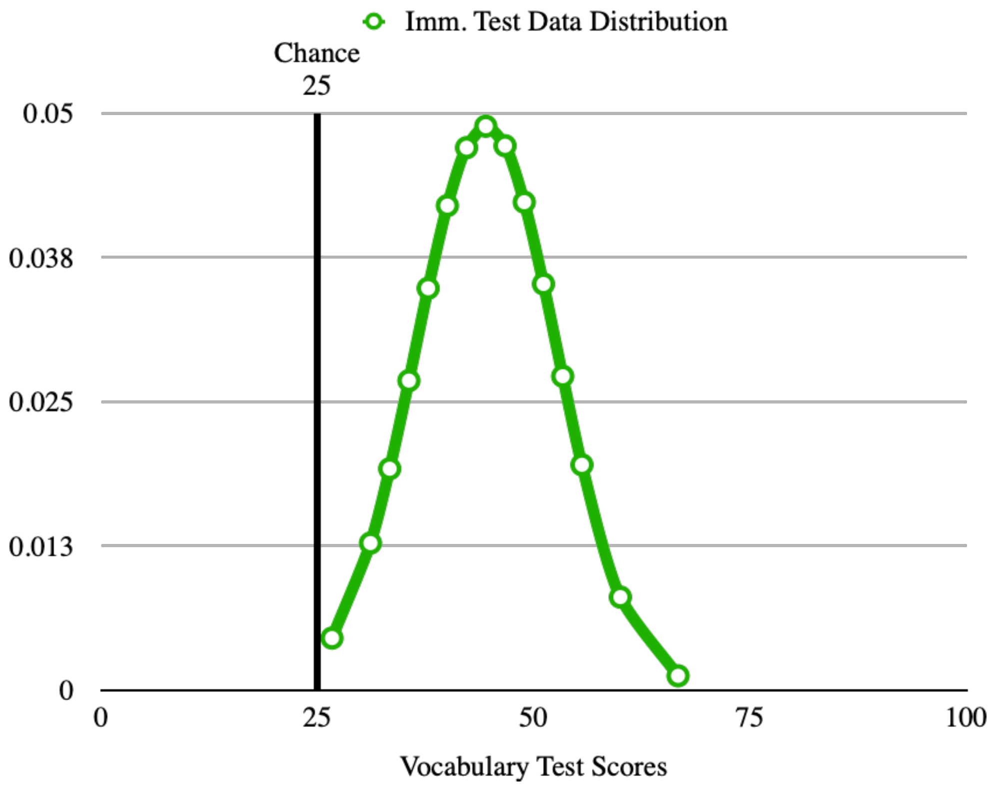


Figure 4.31 Delayed post-test normal data distribution from LC group

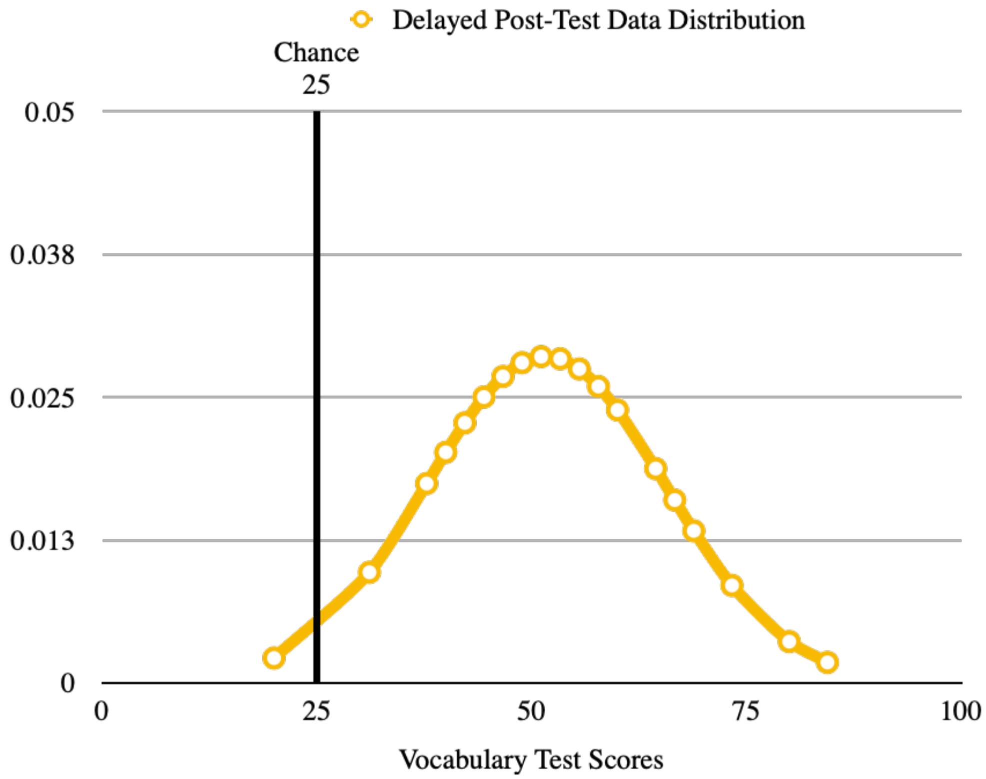


Figure 4.32 OSD post-test normal data distribution from LC group

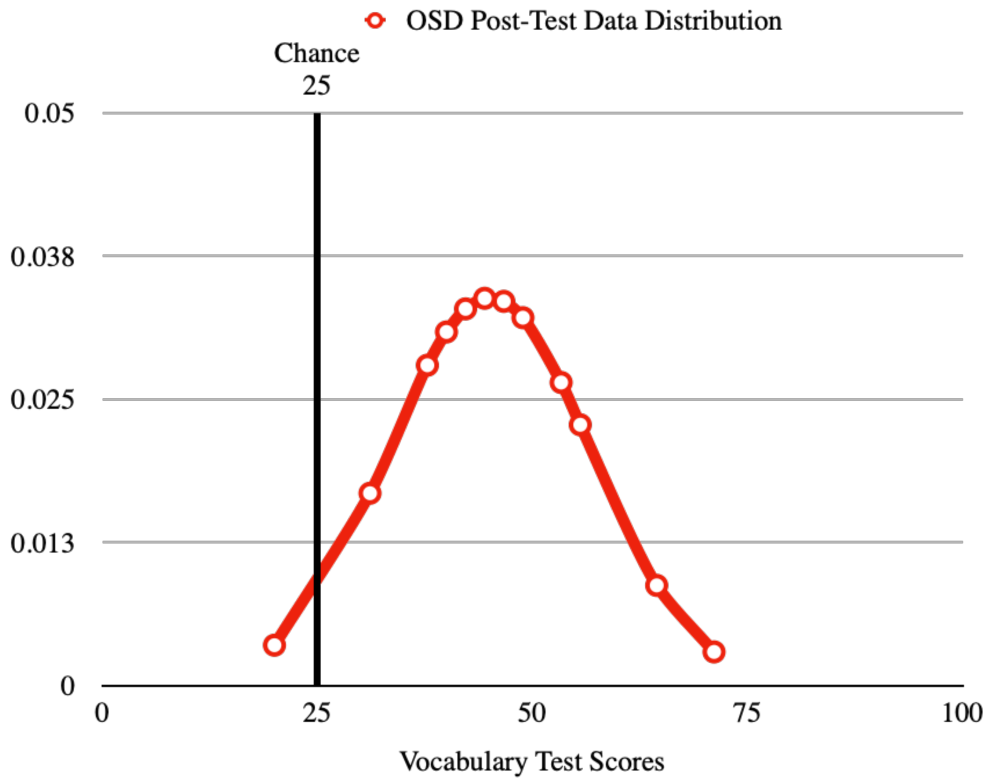


Figure 4.33 Comparative data distribution from LC group (pre-test and delayed post-test values)

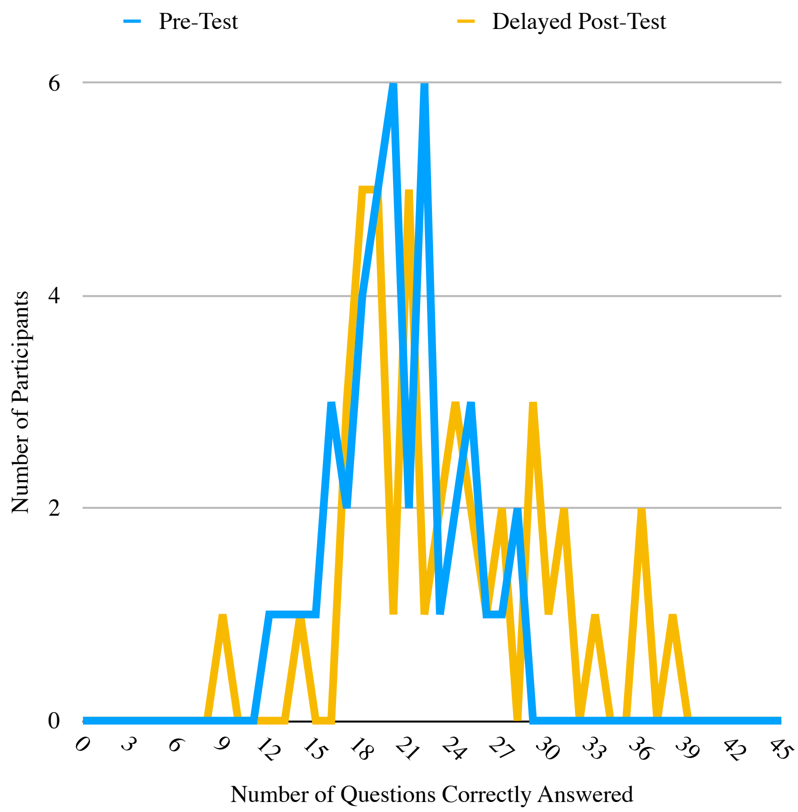


Figure 4.34 Comparative data distribution from LC group (pre-test 1 and delayed post-test 1)

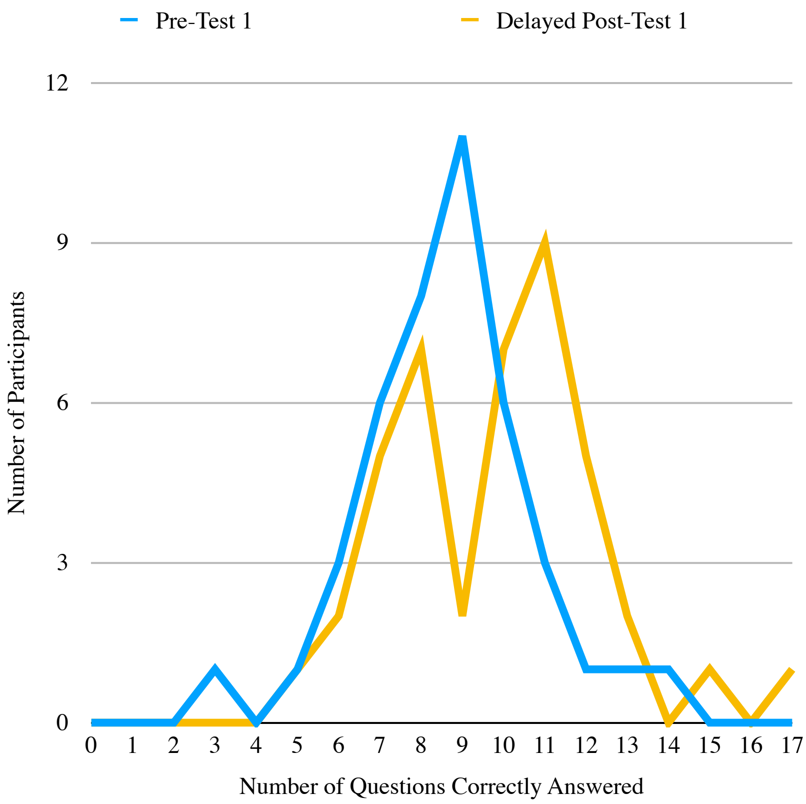


Figure 4.35 Comparative data distribution from LC group (pre-test 3 and delayed post-test 3)

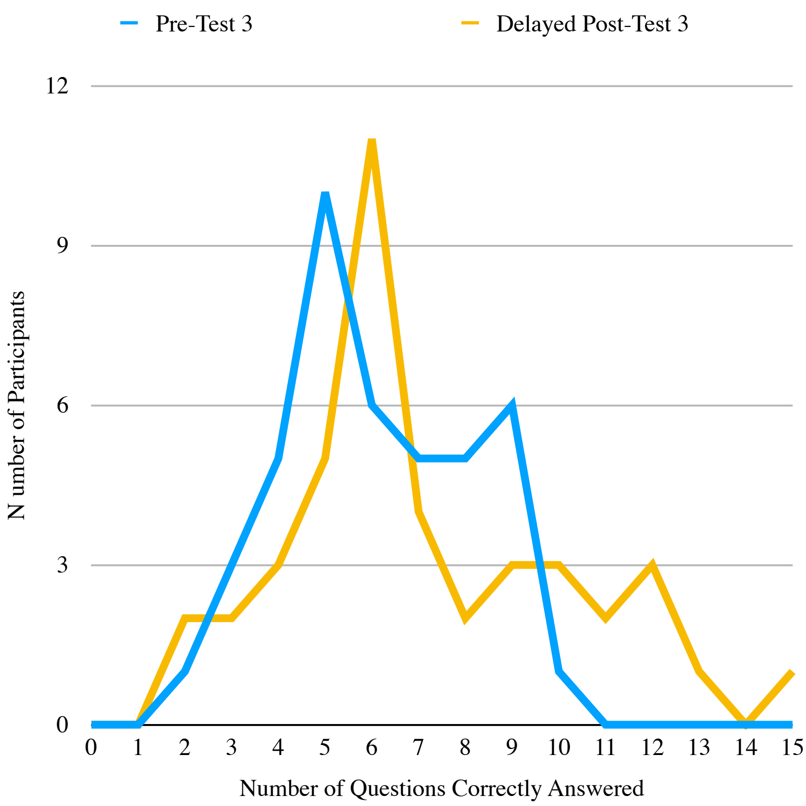


Figure 4.36 Comparative data distribution from LC group (pre-test 3 and immediate test 3)

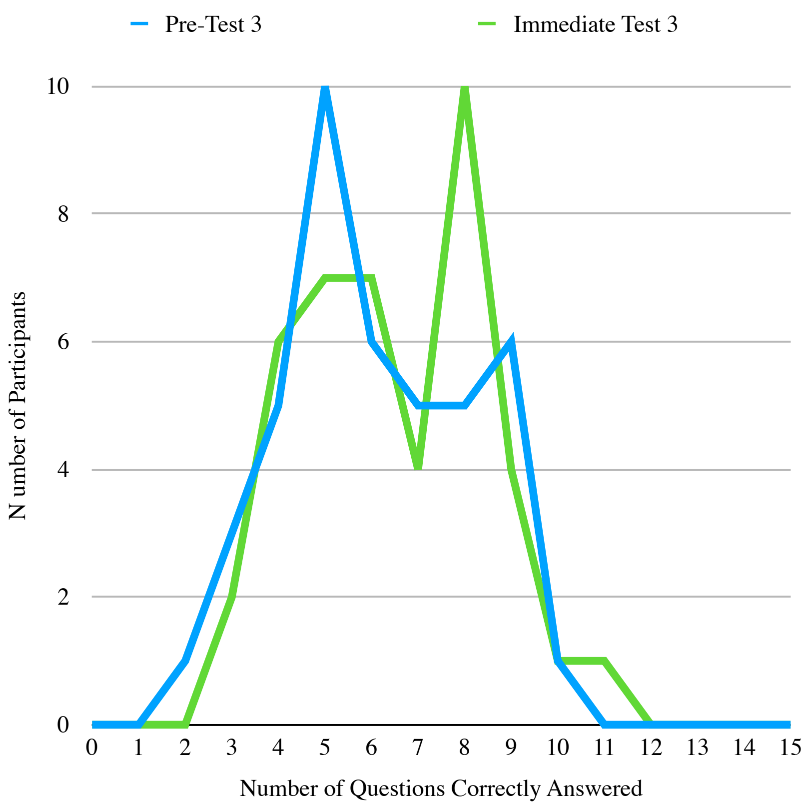


Figure 4.37 Comparative data distribution from LC group (pre-test 2 and OSD post-test 2)

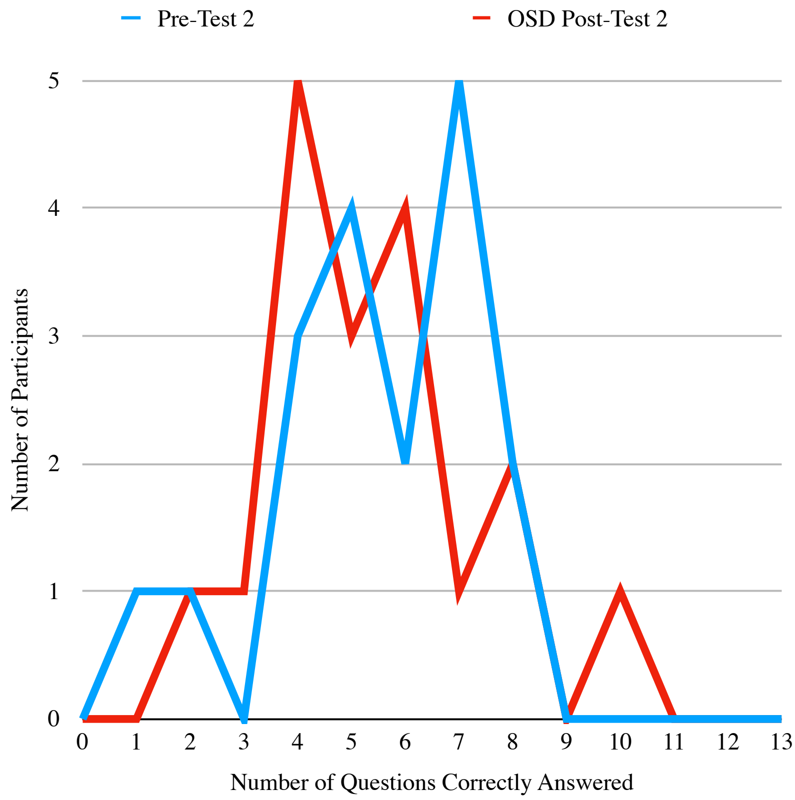
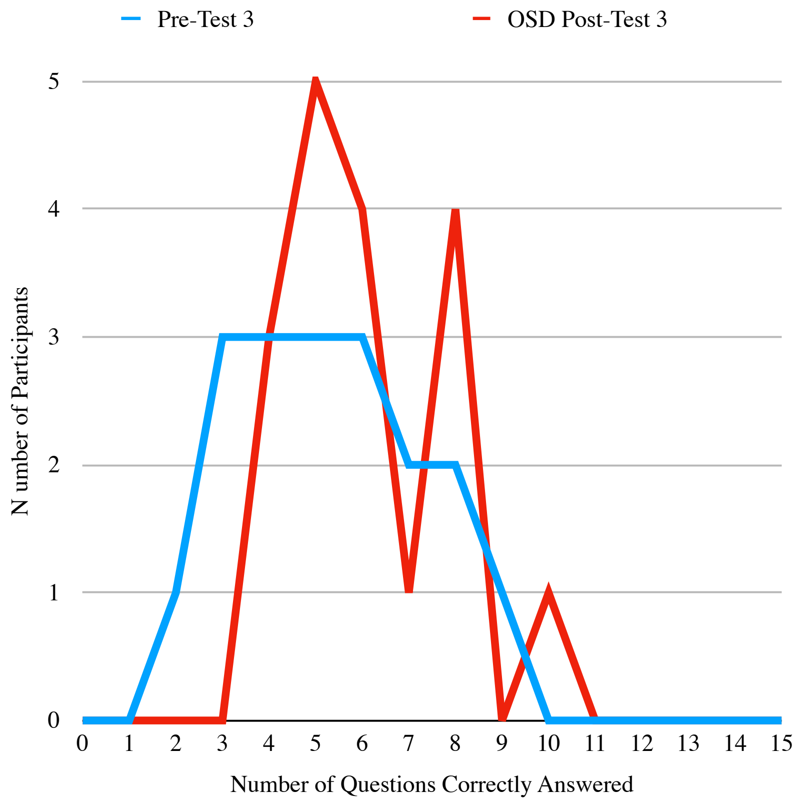


Figure 4.38 Comparative data distribution from LC group (pre-test 3 and OSD post-test 3)



4.8.1 Analysis of non-normal data distribution in the modified LC group sample

The skew values for the delayed post-test protocol in the LC group are mixed between positive and negative skews. The skew value for the total averages (see Figure 4.33) is 0.62 indicating a significant number of participants who are moderately skewing the median and mean higher than the mode of the dataset. The excess kurtosis value of -0.23 is within normal data distribution parameters. This positive skew is mirrored in delayed post-test three (see Figure 4.35) with a skew value of 0.37 (low skew) and an excess kurtosis value of -0.4 (normal range). However, these positive skews are countered by delayed vocabulary post-tests one and two (see Figure 4.34) which both display negative skews of -0.47 and -0.31 (low skew) and excess kurtosis values of 0.19 and 0.042 (normal ranges) respectively. The mix between positive and negative skews in the delayed post-test protocol makes it difficult to generalize any interpretations as to the source of the data asymmetry. However, considering that the only significant change in vocabulary test results occurred in the delayed post-test protocol, it may be argued that some students in the LC sample were more receptive to the corrective feedback protocol for specific vocabulary test items than others.

Another source of negative skew was observed in immediate test three (see Figure 4.36). With a skew value of -0.77, the data distribution may be described as having a moderately strong negative skew with a significant number of participants scoring below the mode of the dataset. Of particular note is the shift in skew from pre-test to immediate test in vocabulary test three. The remainder of the immediate vocabulary tests both display approximate symmetrical data distribution, and the total averages for the immediate test protocol conform to a normal data distribution. It may therefore be argued that vocabulary test three contained more difficult test items for the participants in the LC group and that this difficulty was isolated to vocabulary test three. The remainder of the non-normal data distributions all display positive skew (see Figure 4.36, Figure 4.37, and Figure 4.38).

Vocabulary pre-test three has a positive skew of 0.54 (moderate) and an excess kurtosis value of -1.01 (normal range). OSD post-test two has a skew value of 0.7 (moderately strong skew) and an excess kurtosis value of -0.31 (normal range). OSD post-test three has a skew value of 0.65 (moderate skew) and an excess kurtosis value of -0.75 (normal range). As with the total positive skew observed in the delayed post-test protocol, the positive skews in the OSD post-test datasets may add further weight to the interpretation that the corrective feedback protocol elicits better long-term retention in a small but significant number of participants in the LC group.

4.9 Comparison of modified vocabulary test results from LG and LC groups

In order to analyze the efficacy of the gaming and testing interventions in the LG group relative to the testing-only intervention in the LC group, a comparison analysis was conducted. The vocabulary test averages are directly compared between both groups in each of the testing protocols and across all three vocabulary tests. Between-group, two-sample, two-tailed *t*-tests were also used to look for statistically significant differences between the vocabulary test averages in the two groups. The two-sample *t*-tests were used to compare each testing protocol and vocabulary test from one group to the corresponding testing protocol and vocabulary test from the other (i.e., LG group pre-test one to LC group pre-test one). The results of the comparison analysis as outlined above are summarized in Table 4.9 and Figure 4.39, Figure 4.40, Figure 4.41, and Figure 4.42. The following section will highlight the most important findings for each of the four testing protocols and across each of the three tests as well as provide an interpretation of the data comparison.

Table 4.9 Comparison of modified vocabulary test results from LG and LC groups

	LC Pre-Test	LG Pre-Test	LC Imm. Test	LG Imm. Test	LC Post-Test	LG Post-Test	LC OSD Post-Test	LG OSD Post-Test
Sample Size	*n*=42	*n*=78	*n*=42	*n*=78	*n*=42	*n*=78	*n*=18	*n*=23
Vocab. Test 1	50.85	49.63	50.57	54.91	57.85	60.04	51.64	61.9
Vocab. Test 2	42.68	40.93	38.1	53.56	48.17	53.46	41.46	50.84
Vocab. Test 3	40.64	47.61	43.18	56.24	47.62	58.21	40.75	58.56
Total Averages	45.08	46.44	44.5	54.95	51.65	57.53	45.07	57.59
t-Test for Test 1		0.62		0.14		0.51		0.067
t-Test for Test 2		0.53		0.0000017		0.21		0.11
t-Test for Test 3		0.0052		0.0000039		0.0048		0.00012
t-Test Total Avg.		0.44		0.0000014		0.064		0.0082

Figure 4.39 Comparison of modified vocabulary results from LG and LC groups (pre-test values)

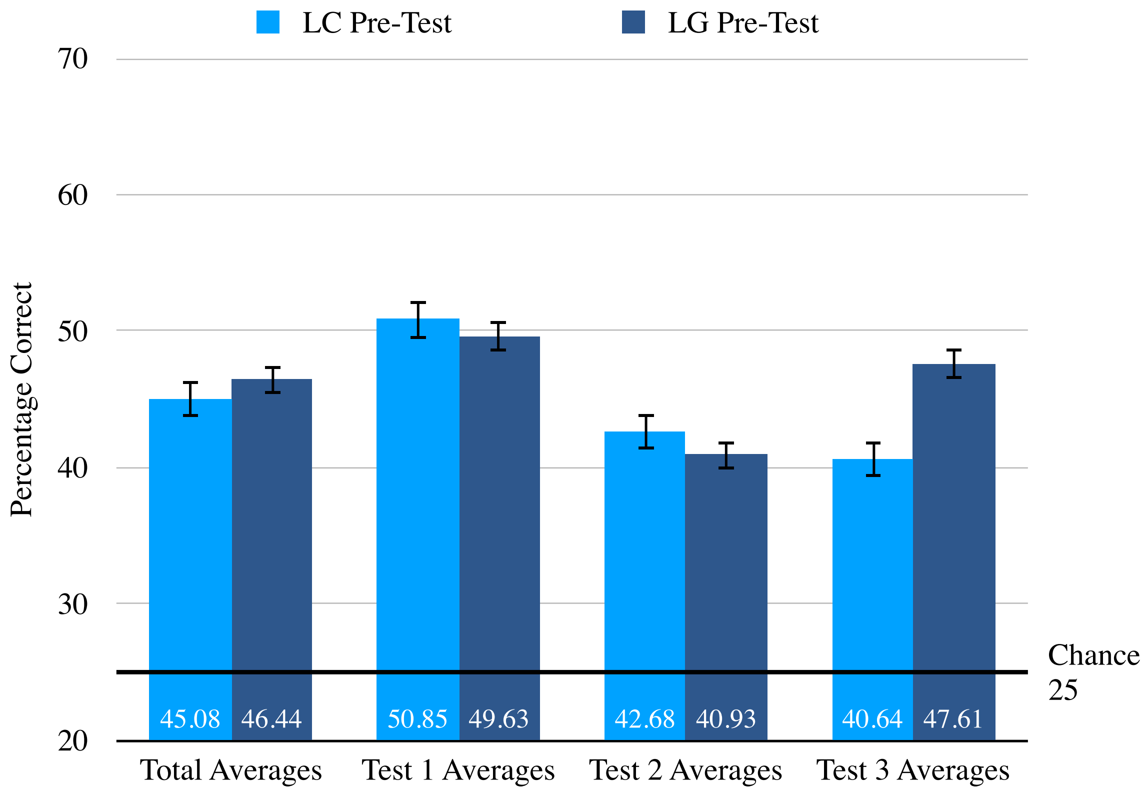


Figure 4.40 Comparison of modified vocabulary results from LG and LC groups (immediate test values)

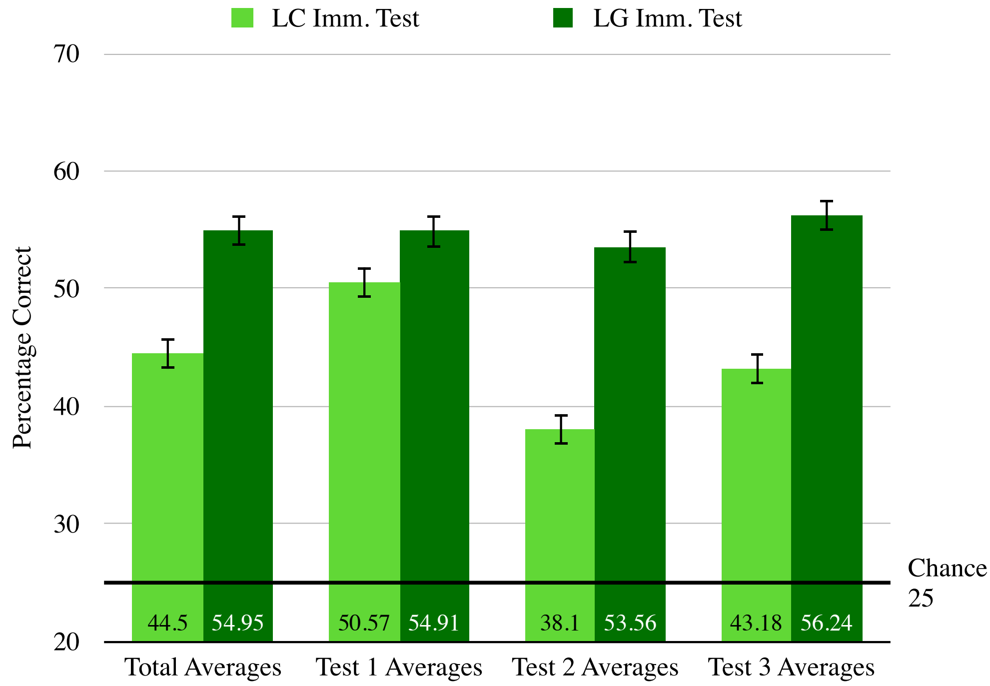


Figure 4.41 Comparison of modified vocabulary results from LG and LC groups (delayed post-test values)

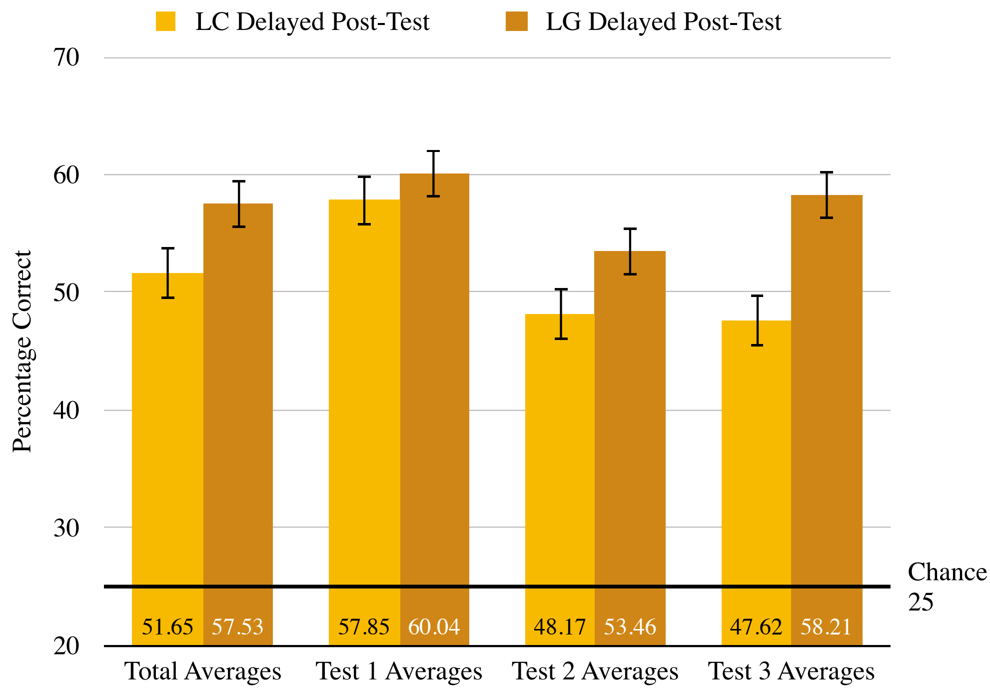
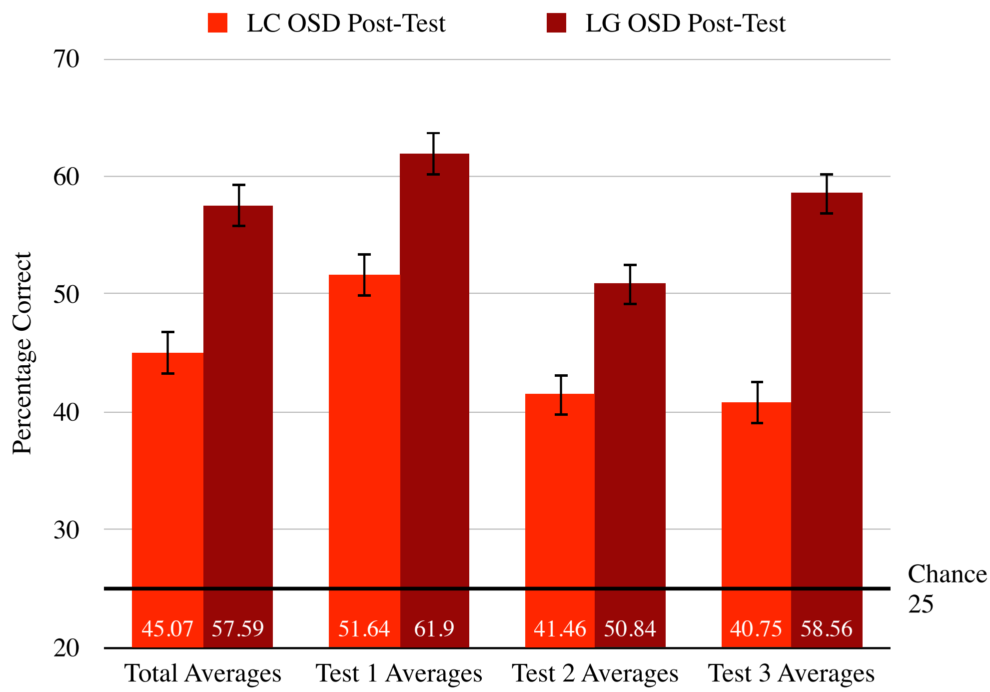


Figure 4.42 Comparison of modified vocabulary results from LG and LC groups (OSD post-test values)



The comparison analysis of the pre-test protocol for both groups revealed that overall both LG and LC groups were statistically similar with no significant differences in vocabulary pre-tests one and two (*p*-values of 0.62 and 0.53 respectively) or in the total averages for the pre-test protocol (*p*-value of 0.44) (see Figure 4.39). With that being said, the vocabulary test averages from vocabulary pre-test three from the LG group were found to be significantly higher than the same test from the LC group with a 6.97 point difference in average test score and a *p*-value of 0.0052 indicating statistically different sample populations. To determine why the results from vocabulary pre-test three were statistically different, the data distribution for each pre-test value were directly compared in an attempt to isolate a potentially skewed data sample as the source of the difference. However, the data distribution from vocabulary pre-test three in the LG group conformed to normal data distribution (see Figure 4.19), whereas the same test in the LC group displayed a statistically significant moderate positive skew (see Figure 4.36). These differences set aside, the total averages for the pre-test protocol in both groups were determined to be statistically similar enough to determine that both groups started their individual study program on approximately equal footing.

With the pre-test values from both groups displaying approximately similar results, the remainder of the testing protocols may be compared to determine the efficacy of the interventions used in each group. Indeed, the immediate test protocol comparison (see Figure 4.40) and OSD post-test comparison (see Figure 4.42) displayed statistically significant differences in the total averages categories (p-values of 0.0000014 and 0.0082 respectively), and the delayed post-test protocol comparison (see Figure 4.41) displayed moderately significant results (p=0.064). These results point strongly to a data interpretation arguing in favor of the overall efficacy of a single-player RPG such as LIS to elicit vocabulary acquisition without specific testing interventions (as is evidenced in the results from the immediate test protocol comparison). Furthermore, these results also argue in favor of the efficacy of the gaming intervention to moderately bolster corrective feedback protocols relative to the LC group sample (as is evidenced in the results from the delayed post-test protocol comparison) and to more effectively elicit long-term vocabulary retention relative to the LC group sample (as is evidenced in the OSD post-test protocol comparison). There were, however, non-significant comparison results when looking at individual vocabulary tests in each of the testing protocols. These results are outlined in detail in the following sections.

4.9.1 Analysis of non-significant findings in the immediate test protocols

In the immediate test protocols, immediate test one was found to be non-significant in the comparison analysis. A comparison of the data distribution of immediate test one from both groups revealed that both samples conformed to normal data distribution, thus making skewed data or high levels of kurtosis implausible explanations as to the source of the anomalous result. However, when looking at the IQA for both groups (which will be detailed in the next chapter, see Figure 5.1 and Figure 5.4), a plausible source for the anomalous results may be found in the statistically significant drops in scores for two test items in the LG group dataset. As will be demonstrated in the IQA analysis, the context of the in-game narrative likely interfered with the selection of the correct Japanese translation for two test items. Although this source is likely not the only reason for the anamolous results observed in the comparison analysis for vocabulary test one, the significant drops in scores for two test items in vocabulary test one for the LG group did arguably shift the total average for vocabulary test one down a couple of points.

4.9.2 Analysis of non-significant findings in the delayed post-test protocols

In the delayed post-test protocols, both delayed post-tests one and two were found to be non-significant in the comparison analysis. A comparison of the data distribution of delayed post-test one and two from both groups revealed that the data from the LC group displayed a low level of negative skew (see Figure 4.34) and that the data from the LG group conformed to normal distribution. However, when a comparison was conducted between the data distribution of delayed post-test three and the total averages for the delayed post-test protocol in both groups, limited evidence of a mismatched comparison were revealed. The data distribution for the delayed post-test protocol in the LG group is slightly negatively skewed (see Figure 4.18) whereas the LC group's total averages are positively skewed to a moderate degree (see Figure 4.33). These results suggest that the total delayed post-test averages in the LG group are slightly lower than would be suggested with normal data distribution and that the LC group's total averages in the same protocol are moderately higher than would be suggested with normal data distribution. This interpretation, of course, does not explain the non-significant results from delayed post-tests one and two in the comparison analysis, but it may suggest that the near-significant results from the total averages comparison may be slightly more significant than what has been indicated in the data should the data have conformed to normal distribution. Overall, however, both groups reacted positively to the corrective feedback protocols and saw significant gains in the vocabulary scores in the delayed post-test protocols.

4.9.3 Analysis of non-significant findings in the OSD post-test protocols

In the OSD post-test protocols, both OSD post-test one and two were found to be non-significant in the comparison analysis. A comparison of the data distribution analyses for both non-significant tests in both groups revealed some differences in skew values which may provide limited evidence as to the source of the anomalous results. The data distribution from OSD post-test one in the LG group is slightly negatively skewed (Figure 4.27), whereas the same test in the LC group was found to conform to normal data distribution. The moderately significant p-value of 0.067 in the comparison analysis between OSD post-test one from both groups would therefore lean slightly more towards significance should the data from the LG group have conformed to a normal data distribution. However, the data distributions from OSD post-test two in both groups were found to be positively skewed to a moderate degree (see Figure 4.5 and Figure 4.37). The data distribution comparison results from OSD post-test two therefore do not offer any plausible explanation for the anomalous results. The relatively smaller sample sizes in the OSD-only samples from both groups may partially explain the non-significant results in the comparison analysis. Overall, as may be argued in the delayed post-test protocols, both groups reacted positively to corrective feedback protocols and displayed some degree of long-term vocabulary retention. However, the LG group displayed significantly more long-term retention in the total averages from the OSD post-test protocol.

4.10 Summary of results for vocabulary acquisition and retention

This chapter presented an analysis of data which directly addresses RQ1: How effective is the LIS gaming intervention at eliciting TL vocabulary acquisition and retention when used in conjunction with pre-determined testing protocols relative to a control group? The data analysis presented in this chapter clearly supports the interpretation that the LIS gaming intervention (i.e., gaming and testing) is 1) more effective at eliciting immediate gains in TL vocabulary acquisition and 2) more effective at eliciting long-term retention of TL vocabulary after corrective feedback relative to the control intervention (i.e., just testing). Furthermore, the approximately symmetrical data distributions in the majority of vocabulary test data from both LG and LC groups also adds weight to the interpretation that these findings may be applicable to demographically similar populations. The following chapter will outline how the LIS gameplay mechanics affected the TL vocabulary acquisition and retention as has been demonstrated in this chapter.

Chapter 5 Results for efficacious gameplay mechanics

The following chapter will detail the results of the data analysis which directly address RQ2: How do the gameplay mechanics in the smartphone application of LIS affect vocabulary acquisition and retention of TL vocabulary? Specifically, the following sections will present the IQA analyses from both LG and LC groups to identify which gameplay mechanics may be implicated in the better TL vocabulary acquisition/retention rates observed in the LG group. Finally, an analysis of the results from the comprehension test protocols from the LG group will be presented to determine whether comprehension of the game's narrative content was correlated to better performance on the vocabulary tests.

5.1 Modified IQA results from LG group

The following section will analyze the number of individual test items for which statistically significant changes in the percentages of correct answers were observed in the LG group. In order to measure statistically significant changes, *t*-tests for individual test items in each testing protocol were used to compare each testing protocol data with vocabulary pre-test results (i.e., immediate tests, delayed post-test, and OSD post-test results). In addition to providing a quantitative analysis of significant shifts in group scores for individual questions for each of the three vocabulary tests, the IQA will also provide an interpretation as to why certain test items underwent significant changes in scores.

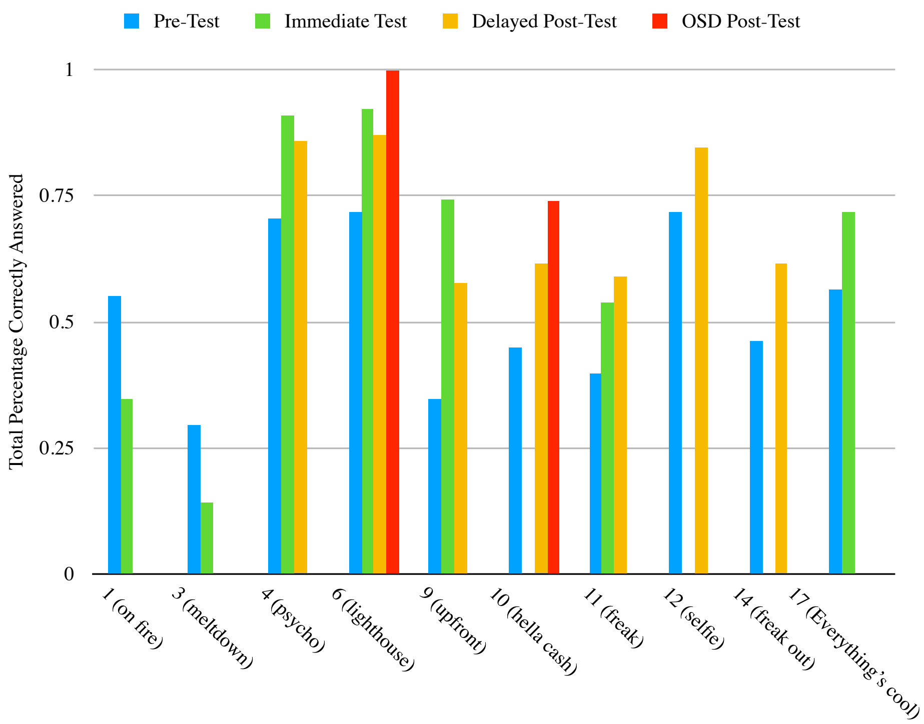
5.1.1 IQA analysis for vocabulary test one (LG group)

There were a total of 17 test items in vocabulary test one, out of which significant changes were observed in 10 of the total test items (58.82%). The test items, percentage changes, and associated testing protocols are summarized in Table 5.1 and Figure 5.1. Of the 10 test items for which significant changes were observed, seven of these items experienced changes in the immediate test protocol. Increases in scores were observed for all test items in the immediate test protocol except for two items, test item one “on fire” and test item three “meltdown,” for which significant drops in the percentages of correctly answered questions were observed. Similarly, gains were observed for seven test items in the delayed post-test protocol. Finally, significant gains in two of the test items were observed in the OSD post-test protocol. The following section will provide an interpretation of the context behind the decreases and increases in scores.

Table 5.1 IQA results from LG group (vocabulary test one)

Test items with significant changes	Pre-test averages	Imm. test averages	Post-test averages	OSD post-test averages
1 (on fire)	0.55	0.35		
3 (meltdown)	0.29	0.14		
4 (psycho)	0.71	0.91	0.86	
6 (lighthouse)	0.72	0.92	0.87	1
9 (upfront)	0.35	0.74	0.58	
10 (hella cash)	0.45		0.62	0.74
11 (freak)	0.4	0.54	0.59	
12 (selfie)	0.72		0.85	
14 (freak out)	0.46		0.62	
17 (Everything’s cool)	0.56	0.72		
Total count	10	7	7	2

Figure 5.1 IQA results from LG group (vocabulary test one)



5.1.2 Analysis of statistically significant decreases in vocabulary test one items (LG group)

The decreases in score on test items one and three in the immediate test protocol may be a result of the in-game narrative content providing misleading contextual information about the definition of each test item. The most common mistaken response for test item one "on fire" was "賢い (clever)." The phrase "to be on fire" in English is commonly used in athletic contexts, whereby a team or individual player is on a winning or scoring streak. The Japanese translation used for the correct answer to this test item in the present study was therefore "連勝中 (winning streak)." However, the context in which the phrase "on fire" was used in the game made the selection of "賢い (clever)" a viable option. During Mr. Jefferson's lecture in the photography class at the beginning of the game, Mr. Jefferson asks Max a number questions in class to which the player must provide the correct answer to. After correctly answering all of Mr. Jefferson's in-class questions, Mr. Jefferson says that Max "is on fire." Given the context of the narrative content in which Max must continually provide correct answers to questions asked by her teacher, the commonly mistaken selection of "賢い (clever)" becomes much more understandable.

The most common mistaken response for test item three "meltdown" was "崩壊 (collapse)." The phrase "meltdown" in English is commonly used when someone is under considerable psychological stress and loses their emotional composure. The Japanese translation used for the correct answer to this test item in the present study was therefore "号泣 (bawling)." However, the context in which the phrase "meltdown" was used in the game made the selection of "崩壊 (collapse)" a viable option. After Max goes to the girls' restroom for the first time after Mr. Jefferson's class, she remarks that she is glad that no one is there to see her "meltdown." In this context, Max is using the phrase "meltdown" hyperbolically, as she is not actually depicted as losing her emotional composure and breaking down into tears.

Given the context of the narrative content in which Max does not actually cry in the girls' bathroom, the much more literal translation "崩壞 (collapse)" of the phrase "meltdown" becomes much more understandable.

5.1.3 Analysis of statistically significant gains in vocabulary test one items (LG group)

Of the 10 total test items in vocabulary test one for which significant changes were observed, five of these maintained significant gains across multiple testing protocols (test items four "psycho," test item six "lighthouse," test item nine "upfront," test item ten "hella cash," and test item eleven "freak"). Of these five test items, three of them (test items four, ten, and eleven) occurred during the action sequence featuring a conflict between the character Chloe and Nathan in the girls' restroom. This conflict scene is one of the few linguistically down-regulated action sequences used in the game, during which the player must manipulate Max with the controls to break open the fire alarm within a short amount of time in order to save Chloe from being shot by Nathan.

Should the player be unable to manipulate Max to break open the fire alarm in time, he/she must rewind the game. While the player attempts to complete the action sequence, the dialogue between Chloe and Nathan continuously plays over and over again in the background. Although there were no data collected on how many times the students repeated the in-game dialogue during the action sequence, the repetition of the in-game dialogue during such a sequence may have played a role in eliciting long-term retention of the vocabulary used both in the dialogue itself and in conjunction therewith after the action sequence took place (e.g., the conversation with principal Wells after the restroom action sequence during which "upfront" is used). Finally, the test item "lighthouse" was the only test item in the entire IQA to experience gains in all testing protocols. These significant gains are arguably due to the multimodal presentation of the text "lighthouse" (i.e., simultaneous presentation of text and voice-acted audio) and numerous visual depictions of the lighthouse in multiple in-game sequences.

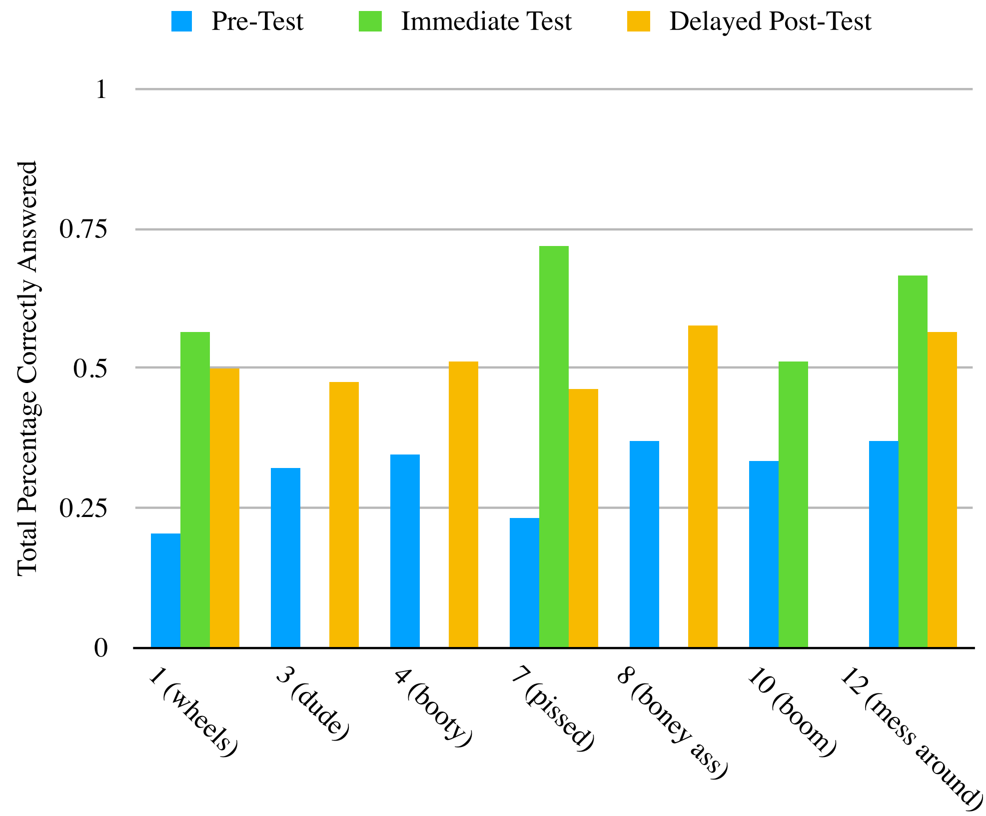
5.1.4 IQA analysis for vocabulary test two (LG group)

There were a total of 13 test items in vocabulary test two, out of which significant changes were observed in seven of the total test items (53.85%). The test items, percentage changes, and associated testing protocols are summarized in Table 5.2 and Figure 5.2. In contrast to the IQA results from vocabulary test one, all significant changes that were observed in vocabulary test two were gains. Of the seven test items for which significant changes were observed, four of these items experienced gains in the immediate test protocol. Gains were observed for six test items in the delayed post-test protocol. Finally, no significant gains were observed in the OSD post-test protocol. The following section will provide an interpretation as to the context behind the increases in scores.

Table 5.2 IQA results from LG group (vocabulary test two)

Test items with significant changes	Pre-test averages	Imm. test averages	Post-test averages	OSD post-test averages
1 (wheels)	0.21	0.56	0.5	
3 (dude)	0.32		0.47	
4 (booty)	0.35		0.51	
7 (pissed)	0.23	0.72	0.46	
8 (boney ass)	0.37		0.58	
10 (boom)	0.33	0.51		
12 (mess around)	0.37	0.67	0.56	
Total count	7	4	6	0

Figure 5.2 IQA results from LG group (vocabulary test two)



5.1.5 Analysis of statistically significant gains in vocabulary test two items (LG group)

Just as with the test item "lighthouse" from vocabulary test one, test item one from vocabulary test two "wheels" is depicted multimodally with Warren talking about his "new wheels" while motioning to his car. For test items seven "pissed" and twelve "mess around," the in-game mechanics arguably helped to elicit the acquisition and retention of these test items. Both test items seven and twelve are associated with the dormitory sequence during which Max must drop paint on Victoria to get her to move out of the way, then provide proof that Victoria was behind the "sexting" incident with Juliet's boyfriend Zach. The game sequences in which both test items appear are both mandatory in-game sub-tasks which require the player complete a series of actions before progressing through the rest of the game. In such required gaming sequences, the context of Victoria being angry (i.e., "pissed") about being splashed with paint and Juliet thinking that Zach is cheating on her (i.e., "messing around") with Dana are arguably better understood by forcing the players to solve narrative puzzles which utilize said contexts before being able to progress any further in the game.

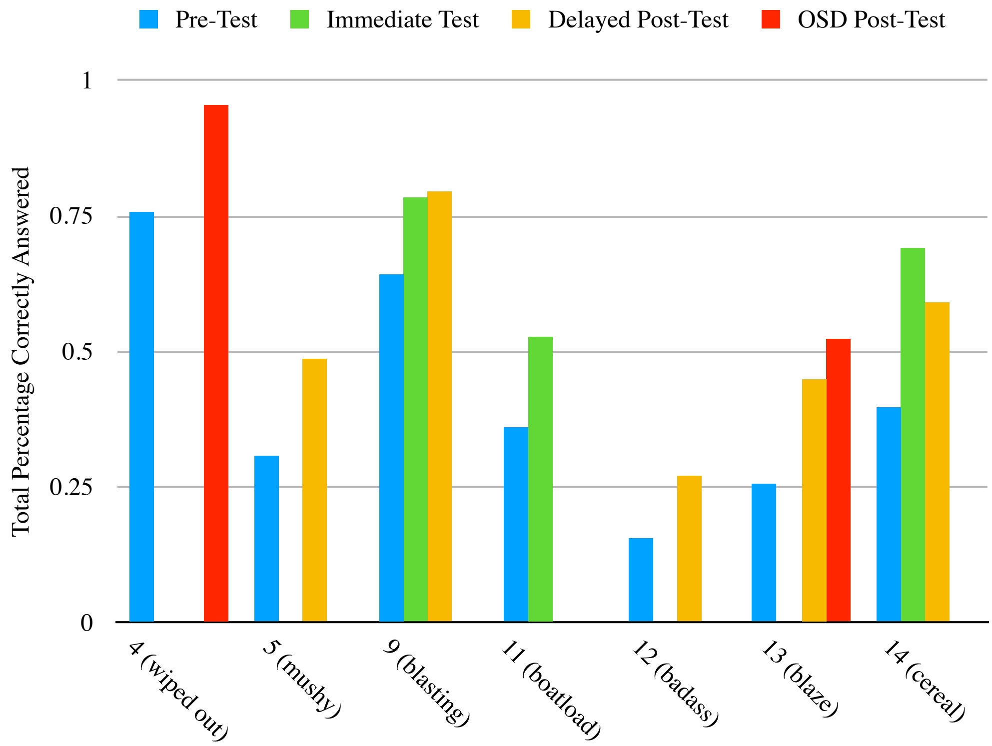
5.1.6 IQA analysis for vocabulary test three (LG group)

There were a total of 15 test items in vocabulary test three, out of which significant changes were observed in seven of the total test items (46.67%). The test items, percentage changes, and associated testing protocols are summarized in Table 5.3 and Figure 5.3. In contrast to the IQA results from vocabulary test one, all significant changes that were observed in vocabulary test three were gains. Of the seven test items for which significant changes were observed, three of these items experienced gains in the immediate test protocol. Gains were observed for five test items in the delayed post-test protocol. Finally, two significant gains were observed in the OSD post-test protocol. The following section will provide an interpretation as to the context behind the increases in scores.

Table 5.3 IQA results from LG group (vocabulary test three)

Test items with significant changes	Pre-test averages	Imm. test averages	Post-test averages	OSD post-test averages
4 (wiped out)	0.76			0.96
5 (mushy)	0.31		0.49	
9 (blasting)	0.64	0.78	0.79	
11 (boatload)	0.36	0.53		
12 (badass)	0.15		0.27	
13 (blaze)	0.26		0.45	0.52
14 (cereal)	0.4	0.69	0.59	
Total count	7	3	5	2

Figure 5.3 IQA results from LG group (vocabulary test three)



5.1.7 Analysis of statistically significant gains in vocabulary test three items (LG Group)

When isolating the test items which experienced gains across multiple testing protocols, hence displaying long-term retention, test item nine "blasting," test item thirteen "blaze," and test item fourteen "cereal" may provide further insight into which test items the gaming and testing interventions elicited significant gains. In vocabulary test one, test items associated with the linguistically down-regulated action sequence in the girls' restroom, during which Max saves Chloe from being shot by Nathan, experienced better long-term retention relative to other vocabulary for which significant changes were observed. The test items "lighthouse" from vocabulary test one and "wheels" from vocabulary test two arguably benefit from the multimodal presentation of the vocabulary. Finally, vocabulary associated with the context behind mandatory in-game tasks such as "pissed" and "mess around" from vocabulary test two also experienced better long-term retention relative to other test items. Similarly, the three test items from vocabulary test three for which significant gains were observed across multiple testing protocols conform to the same patterns conducive to long-term retention as described above.

Test item fourteen “cereal” is a nonstandard morph of the word “serious” and is used by Max to highlight her disappointment that her camera broke, as in, “Are you serious?!” Although there is no direct visual representation of the word “cereal,” there are numerous visual depictions of Max attempting to fix her camera after it is broken. Furthermore, there is a mandatory in-game task devoted to finding the tools with which Max attempts to fix said broken camera. The combination of the multimodal presentation of the broken camera with which the phrase “cereal” is associated and the context behind the mandatory in-game task during which Max searches for the tools to try and fix her broken camera arguably work synergistically to elicit long-term retention of the test item.

Similarly, test item thirteen “blaze” is a nonstandard word meaning “to smoke marijuana” which is depicted visually with Chloe smoking a joint on her bed. The context of the consumption of marijuana is further elaborated upon with the linguistically down-regulated action sequence during which Max must hide from Chole’s stepfather, David. The subsequent dialogue between Chloe and David mentions the use of marijuana on numerous occasions, for which the player must make the in-game narrative choice to take the blame for the marijuana or to deny the use and implicate Chloe.

Finally, during the in-game action sequence during which Max must hide from David, test item nine “blasting” is mentioned (potentially multiple times depending on how many times the player requires to complete the action sequence to hide in the closet before David comes into the room) when David complains that Chloe is “blasting” music. Indeed, the playing of the very same music which David complains that Chloe is “blasting” is subject to the multimodal presentation of the girls dancing to the music and a mandatory in-game sub-task, whereby the player must locate and play a CD in Chloe's stereo before being able to progress any further in the game.

5.1.8 Gameplay mechanics implicated in statistically significant gains

There were significant gains observed throughout all three testing protocols in the IQA. Although overall there were more gains observed in the delayed post-test protocol after receiving corrective feedback, the gains observed in the immediate test protocols were often higher than in the subsequent delayed post-test protocol. Test items four "psycho," six "lighthouse," and nine "upfront" from vocabulary test one, test items one "wheels," seven "pissed," and twelve "mess around" from vocabulary test two, and test item fourteen "cereal" from vocabulary test three displayed higher percentages of correct answers in the immediate test protocols relative to delayed post-test protocols. Therefore, the corrective feedback protocol is undeniably eliciting significant gains in vocabulary acquisition and retention. However, the gaming intervention itself also undeniably plays a role in eliciting vocabulary acquisition and retention. The game mechanics which seem particularly efficacious at eliciting vocabulary acquisition and retention include linguistically down-regulated action sequences, mandatory in-game tasks, and multimodal presentation of vocabulary. Furthermore, as will be discussed in detail in the analysis of the comprehension test results, the LG-group IQA results provide evidence to support the interpretation that these gameplay mechanics may elicit higher levels of narrative comprehension, which in turn resulted in better performance for the vocabulary test items described above.

5.2 Modified IQA results from the LC group

The following section will analyze the number of individual test items for which statistically significant changes in the percentages of correct answers were observed in the LC group. The IQA for the LC group was conducted under the same parameters as outlined in the IQA for the LG group. The following sections will detail the test items for which significant changes were observed for each testing protocol across each of the three vocabulary tests. Furthermore, an interpretation is provided as to the source of the significant changes in the LC group.

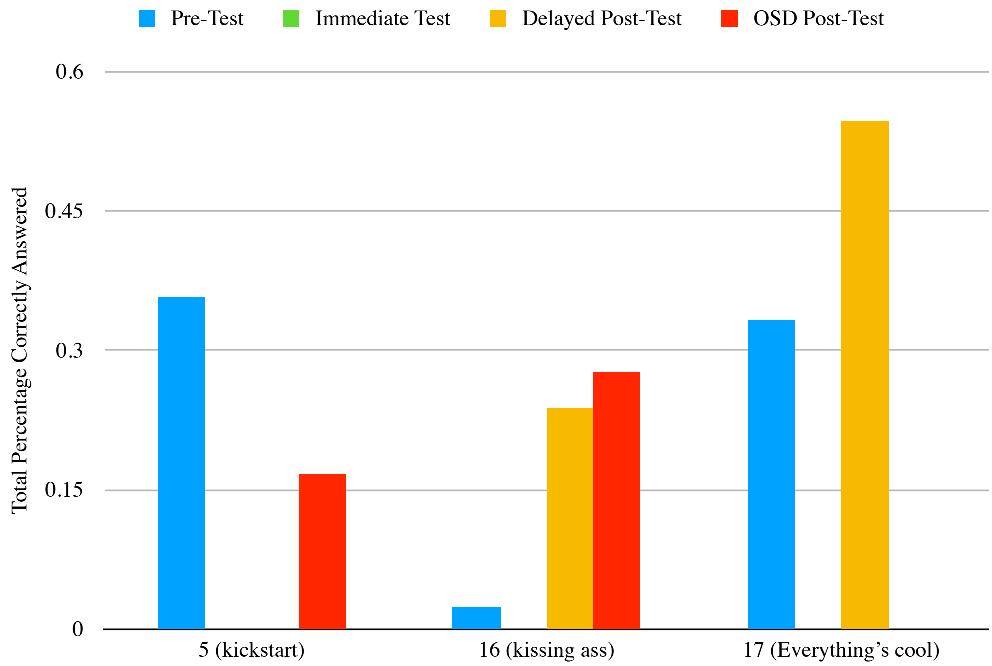
5.2.1 IQA analysis for vocabulary test one (LC group)

There were a total of 17 test items in vocabulary test one, out of which significant changes were observed in three of the total test items (17.65%). The test items, percentage changes, and associated testing protocols are summarized in Table 5.4 and Figure 5.4. As the data show, there were no significant changes observed in the immediate test protocol. These results conform to the vocabulary test results from the LC group. There were two significant gains observed in test item sixteen “kissing ass” and seventeen “everything’s cool” from the delayed post-test protocol. In the OSD post-test protocol, there was one significant drop in score observed in test item five “kickstart” and one significant gain in test item sixteen “kissing ass.” The only test item throughout the entire IQA for the LC group which saw significant growth in multiple testing protocols is test item sixteen “kissing ass.” There was no significant change observed for this specific test item in the IQA for the LG group. Although any interpretation as to why this specific test item spurred long-term retention would be entirely speculative, as all observed vocabulary acquisition occurred after the presentation of corrective feedback, there were arguably a number of students who took notice of the corrective feedback for this specific test item more than others.

Table 5.4 IQA results from LC group (vocabulary test one)

Test items with significant changes	Pre-test averages	Imm. test averages	Post-test averages	OSD post-test averages
5 (kickstart)	0.36			0.17
16 (kissing ass)	0.02		0.24	0.28
17 (Everything’s cool)	0.33		0.55	
Total count	3	0	2	2

Figure 5.4 IQA results from LC group (vocabulary test one)



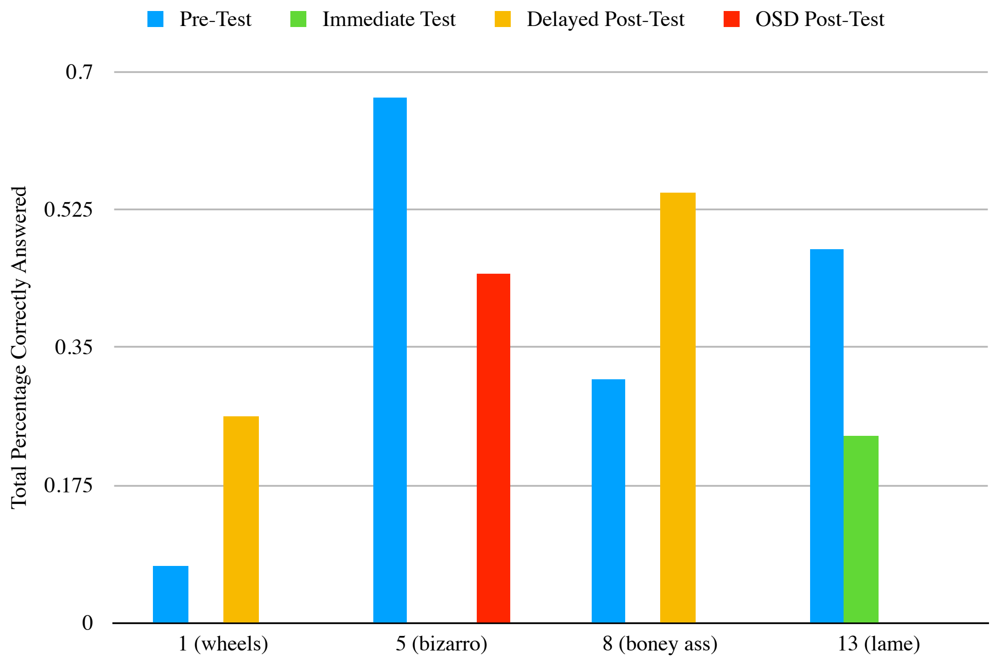
5.2.2 IQA analysis for vocabulary test two (LC Group)

There were a total of 13 test items in vocabulary test two, out of which significant changes were observed in four of the total test items (30.77%). The test items, percentage changes, and associated testing protocols are summarized in Table 5.5 and Figure 5.5. In contrast to the IQA results from vocabulary test one in the LC group, the IQA results for vocabulary test two revealed a significant change in the immediate test protocol, for which a drop in score for test item thirteen "lame" was observed. There were two significant gains observed in the delayed post-test protocol for test items one "wheels" and test item eight "boney ass." Finally, in the OSD post-test protocol there was a significant drop in score for test item five "bizarro." The mixed results of significant gains and drops in scores make it difficult to interpret the source of these significant changes.

Table 5.5 IQA results from LC group (vocabulary test two)

Test items with significant changes	Pre-test averages	Imm. test averages	Post-test averages	OSD post-test averages
1 (wheels)	0.071		0.26	
5 (bizarro)	0.67			0.44
8 (boney ass)	0.31		0.55	
13 (lame)	0.48	0.24		
Total count	4	1	2	1

Figure 5.5 IQA results from LC group (vocabulary test two)



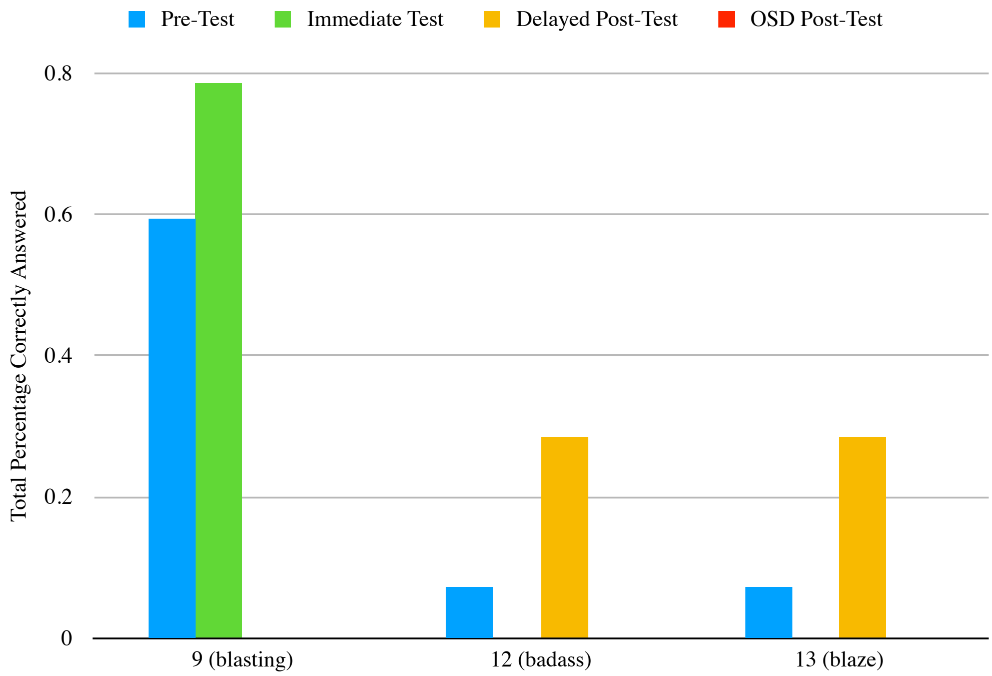
5.2.3 IQA analysis for vocabulary test three (LC Group)

There were a total of 15 test items in vocabulary test three, out of which significant changes were observed in three of the total test items (20%). The test items, percentage changes, and associated testing protocols are summarized in Table 5.6 and Figure 5.6. In contrast to IQA results from both vocabulary test one and two in the LC group, the IQA results from vocabulary test three reveal a significant gain in the immediate test protocol with test item nine “blasting.” The isolated nature of this one significant gain in the immediate test protocol makes it difficult to interpret the source of the gain. There were also two significant gains observed in the delayed post-test protocol with test items twelve "badass" and thirteen "blaze." In line with the previous interpretations for the gains in the delayed post-test protocol in the LC group, the corrective feedback protocol would arguably be the primary source for the gains in score. Finally, there were no observed changes of statistical significance observed in the OSD post-test protocol from vocabulary test three.

Table 5.6 IQA results from LC group (vocabulary test three)

Test items with significant changes	Pre-test averages	Imm. test averages	Post-test averages	OSD post-test averages
9 (blasting)	0.6	0.79		
12 (badass)	0.071		0.29	
13 (blaze)	0.071		0.29	
Total count	3	1	2	0

Figure 5.6 IQA results from LC group (vocabulary test three)



5.3 Comparison of modified IQA results from the LG and LC groups

To add further weight to the strong evidence for the efficacy of the gaming and testing interventions in the LG group over a testing-only intervention in the LC group as outlined above, the following section will outline a comparison analysis of the IQA results from both groups. The total counts for significant increases and decreases in correctly answered percentages of test items from each testing protocol and vocabulary test in both groups are summarized in Table 5.7. As is evidenced in Table 5.7, the LG group saw a total of 34 statistically significant increases in score throughout the study program and only two instances of statistically significant decreases in score. In contrast to these results, the LC group saw only eight statistically significant increases in score and three instances of statistically significant decreases in score throughout the study program. Of further note is the significant difference between score increases in the immediate test protocol at twelve to one. These results add further evidence to the efficacy of single-player RPGs such as LIS to not only increase vocabulary acquisition of in-game vocabulary without the aid of testing interventions but also to bolster corrective feedback protocols used in testing procedures.

Table 5.7 Comparison of IQA results from LG and LC groups

Significant changes	LG imm. test	LC imm. test	LG post-test	LC post-test	LG OSD post-test	LC OSD post-test
V. test 1 increase	5	0	7	2	2	1
V. test 1 decrease	2	0	0	0	0	1
V. test 2 increase	4	0	6	2	0	0
V. test 2 decrease	0	1	0	0	0	1
V. test 3 increase	3	1	5	2	2	0
V. test 3 decrease	0	0	0	0	0	0
Total increase	12	1	18	6	4	1
Total decrease	2	1	0	0	0	2

5.4 Modified comprehension test results

In addition to immediate vocabulary tests one, two, and three, the LG group was also required to take comprehension tests A, B, and C on the same testing schedule. The LC group was not required to complete these comprehension tests as the test items were based solely on the in-game narrative content from LIS. As previously described, the comprehension tests were conducted entirely in English and were designed to 1) measure the students' level of comprehension of the in-game narrative content, and 2) act as a baseline for the students' level of English proficiency.

The results of the comprehension tests are summarized in Table 5.8. In addition to analyzing the results of the comprehension tests, a gender comparison of the total comprehension and vocabulary test averages was carried out to look for significant in-group differences in test scores (see Figure 5.7 and Figure 5.8). An average of total immediate test scores and delayed post-test scores was used for all comparison analyses. Finally, in order to look for a correlation between in-game narrative content comprehension/general English proficiency and vocabulary test scores, correlation analyses were conducted on the aggregate, male, and female samples (see Figure 5.9, Figure 5.10, and Figure 5.11).

Table 5.8 Modified comprehension test results

	Score averages	Correlation to vocabulary tests (*r*-value)	*p*-value for correlation coefficients
Sample size	*n*=42	*n*=42	*n*=42
Comprehension test A	66.83	0.33	0.033
Comprehension test B	62.06	0.41	0.0077
Comprehension test C	59.19	0.19	0.24
Total averages	62.86	0.6	0.000027

Figure 5.7 Gender comparisons of comprehension test averages from LG group

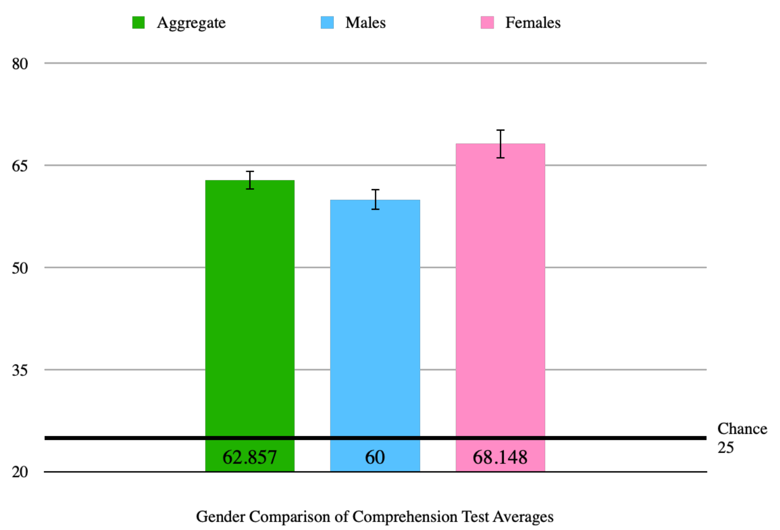


Figure 5.8 Gender comparisons of vocabulary test averages from LG group

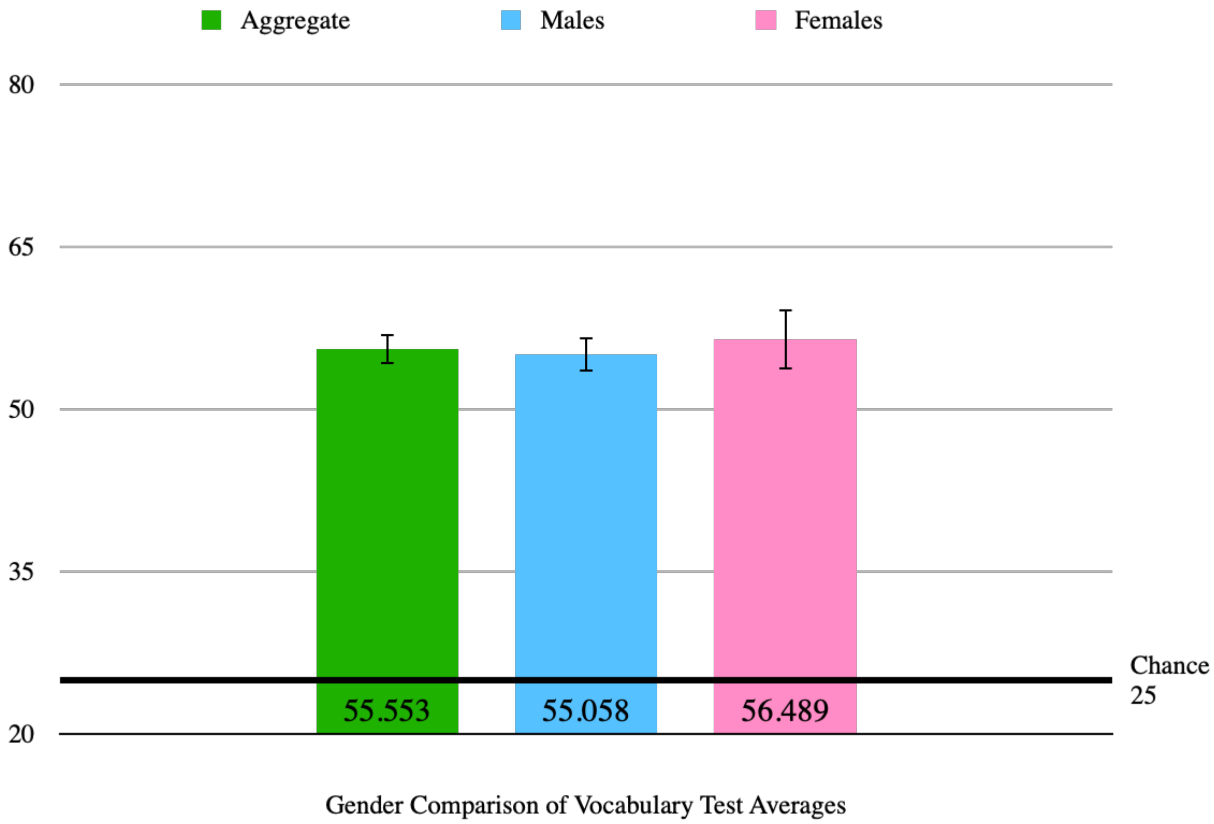


Figure 5.9 Correlation analysis for total comprehension and vocabulary test averages (aggregate sample)

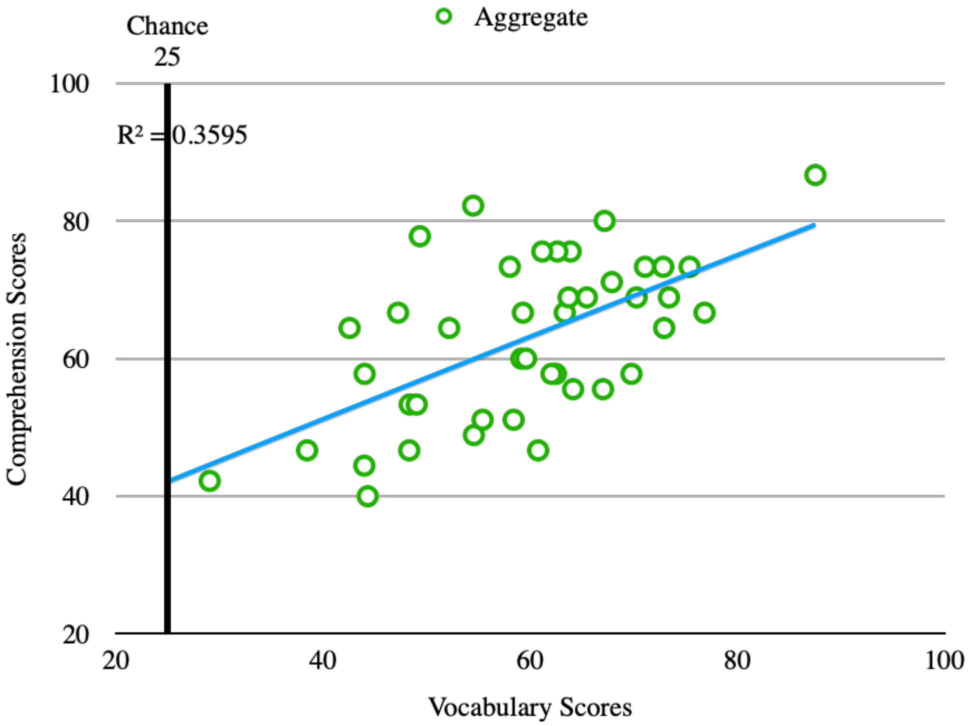


Figure 5.10 Correlation analysis for total comprehension and vocabulary test averages (male sample)

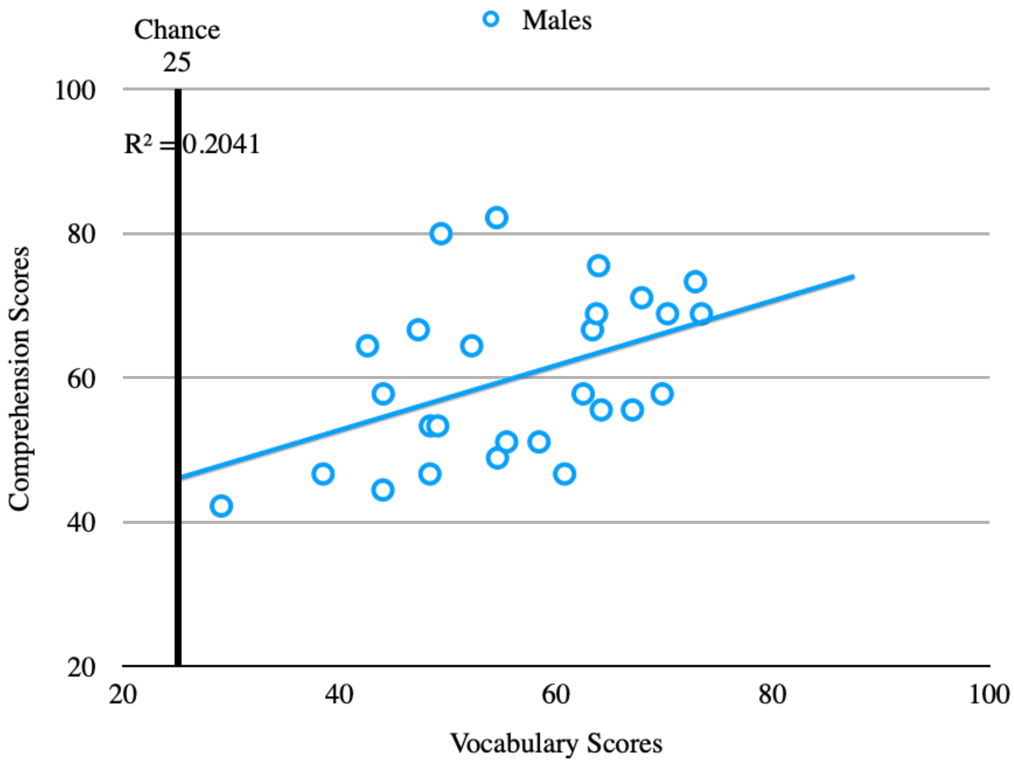
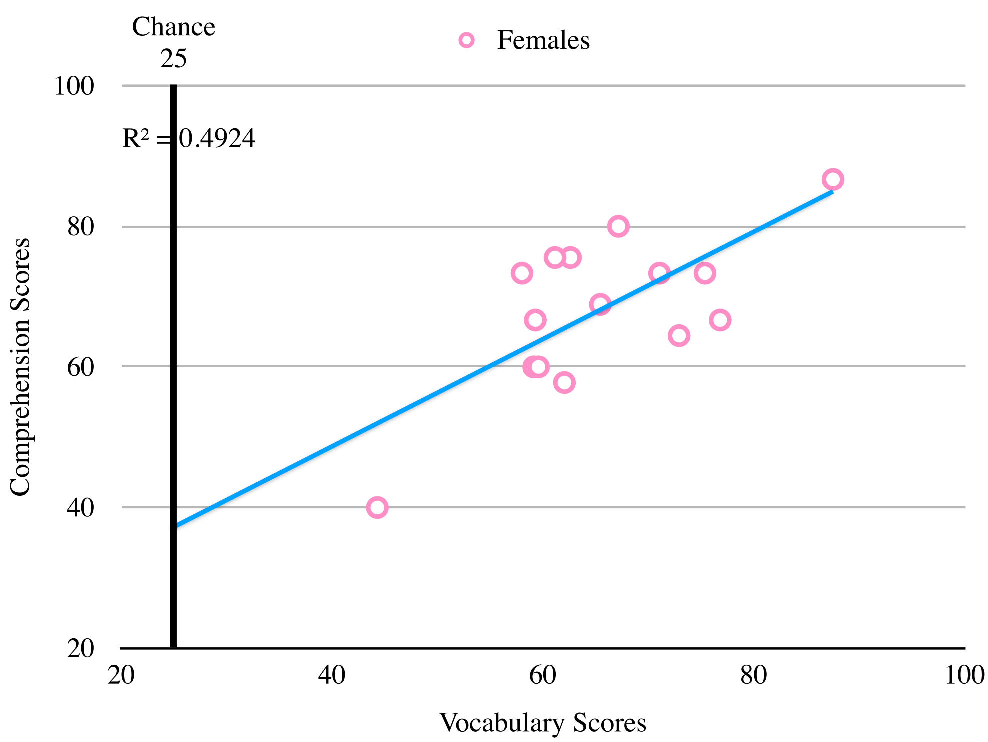


Figure 5.11 Correlation analysis for total comprehension and vocabulary test averages (female sample)



The total comprehension test averages for the female group were observed to be significantly higher than the male group (p=0.03) (see Figure 5.7). There were no statistically significant differences between gender groups for the total vocabulary test averages (p=0.63) (see Figure 5.8). When analyzing the correlation between total comprehension and vocabulary test averages, a strong positive correlation was observed (r=0.6, p=0.000027) potentially indicating that the better the students understood the in-game narrative content (i.e., through corrective gameplay mechanics) or otherwise the better the students' English proficiency level, the higher they were able to score on the vocabulary tests (see Figure 5.9). Alternatively, the opposite data interpretation is possible, whereby the better the students understood the vocabulary test items, the better they understood the in-game narrative content. The strong correlations observed between the comprehension and vocabulary tests were stronger for the female group (r=0.7, p=0.0036) (see Figure 5.11) than the male group (r=0.45, p=0.018) (see Figure 5.10). The difference between the gender group r-values in the correlation analysis described above was also found to be close to statistical significance after transforming the r-values into z-values with Fisher's z transformation and then performing an observed z calculation for both gender values. The observed z-value was 1.54, correlating to a p-value of 0.062. These gender differences potentially indicate that comprehension of in-game narrative content was more of a predictive factor for higher vocabulary test scores in the female group than the male group.

5.5 Analysis of comprehension data distribution

In order to establish the applicability of the comprehension test scores to the vocabulary test scores in the LG group, an analysis of the data distribution of the comprehension test was conducted. Determining normal data distribution was analyzed by calculating skew, excess kurtosis, and their respective z-values. The results of the data distribution analysis for the comprehension tests are summarized in (Table 5.9). The total averages for the comprehension tests and comprehension test C were determined to conform to normal data distribution ranges (see Figure 5.12 and Figure 5.13). However, comprehension tests A and B were both found to be outside of normal data distribution ranges, albeit approximately symmetrical.

Table 5.9 Analysis of comprehension data distribution

	Standard deviation	*z*-value for skew	*z*-value for kurtosis
Sample size	n=42	n=42	n=42
Comprehension test A	15.12	-2.79	-3.81
Comprehension test B	19.45	-4.87	-1.62
Comprehension test C	13.6	-0.38	-0.19
Total averages	11.7	-2.11	-1.61

Figure 5.12 Normal data distribution for total comprehension averages

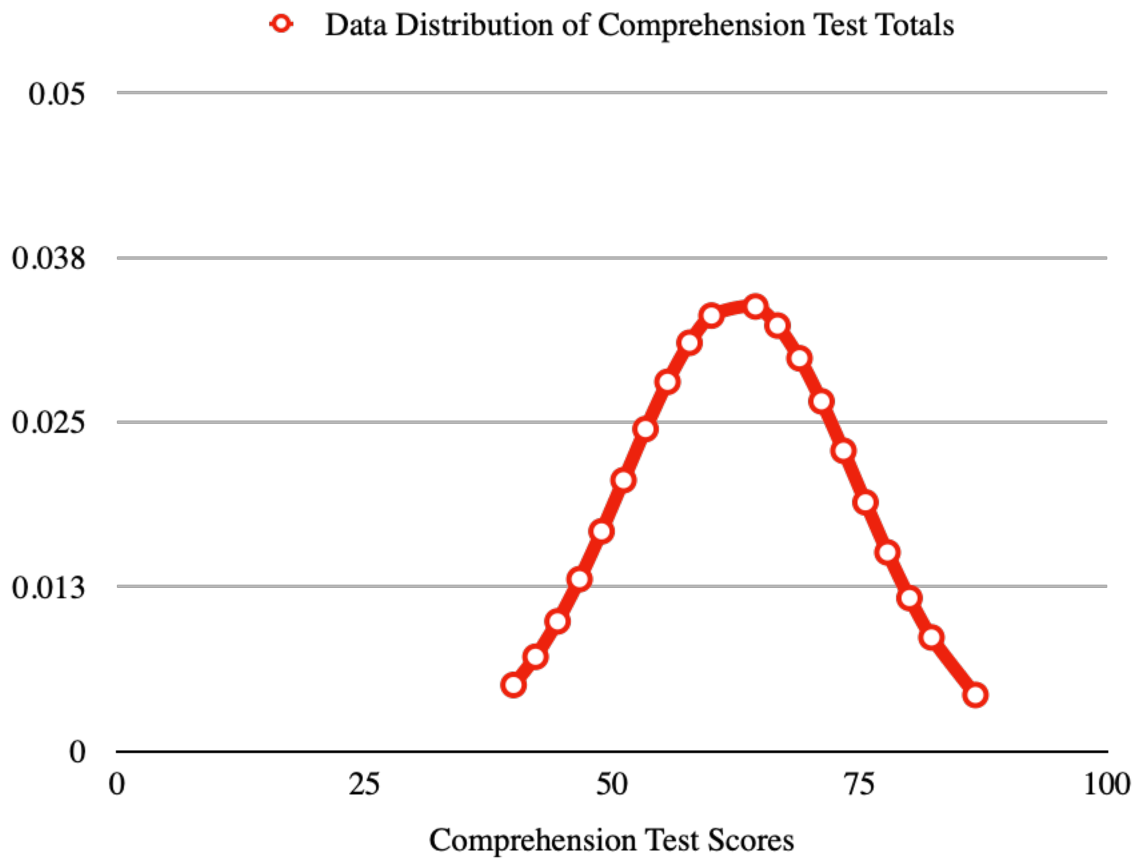
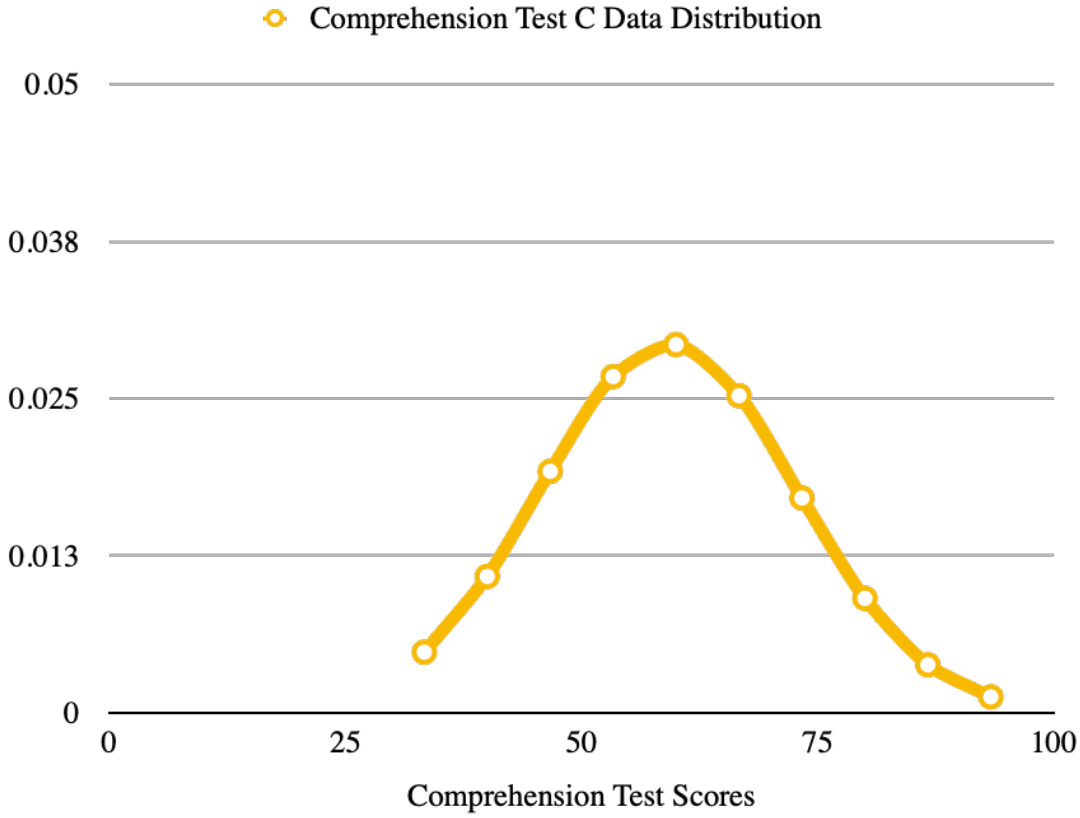


Figure 5.13 Normal data distribution for comprehension test C



The skew value for comprehension test A (-0.43) correlated to approximately symmetrical data distribution, however the excess kurtosis value of -1.35 indicates a leptokurtic distribution of the outliers, whereby the sample population is concentrated around the mean of the dataset (see Figure 5.14). In contrast to these findings, the excess kurtosis value from comprehension test B of -1.02 correlated to normal data distribution ranges, however the skew value of -0.75 indicated a strong negative skew (see Figure 5.15). Overall, the total averages for the comprehension tests are within normal data distribution ranges and are therefore suitable for correlational analyses to the normally distributed vocabulary test averages observed in the LG group.

Figure 5.14 Data distribution for comprehension test A

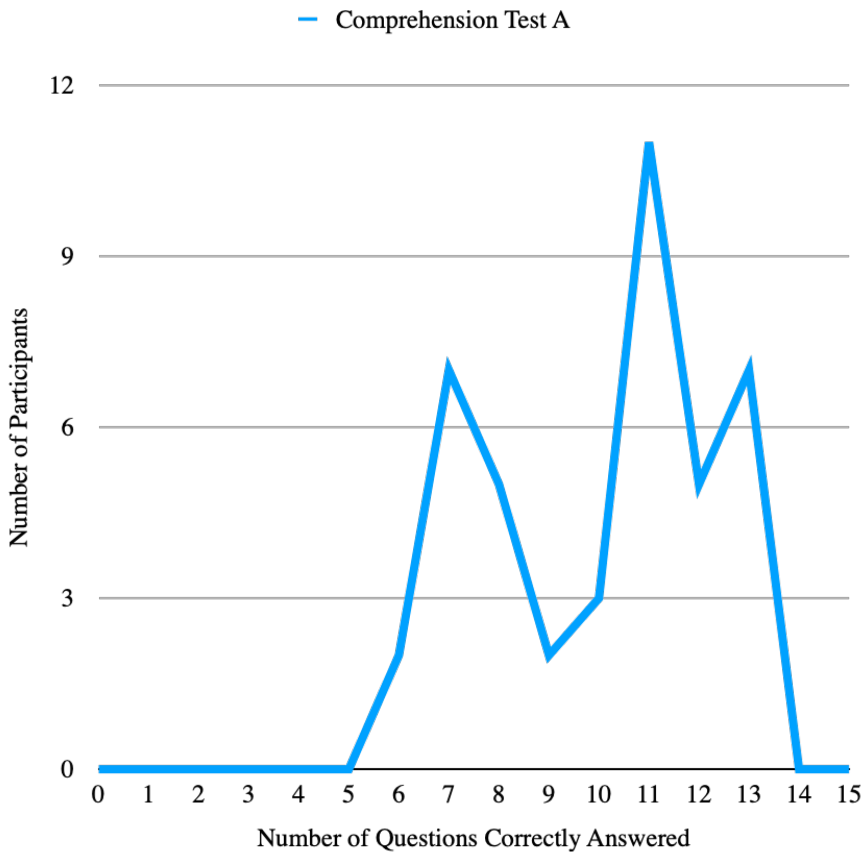
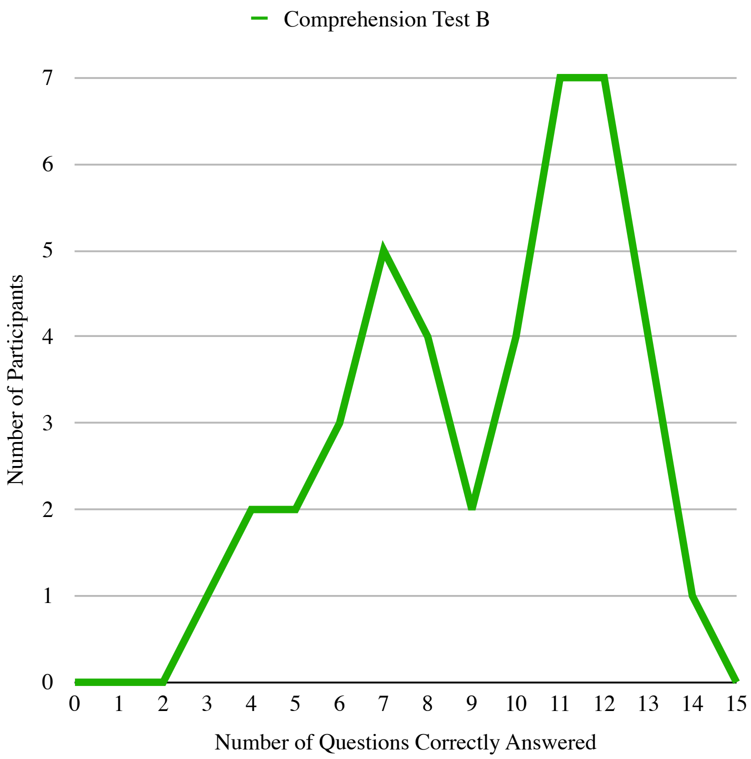


Figure 5.15 Data distribution for comprehension test B



5.6 Summary of results for efficacious gameplay mechanics

This chapter presented an analysis of data which directly addresses RQ2: How do the gameplay mechanics in the smartphone application of LIS affect vocabulary acquisition and retention of TL vocabulary? The data analysis presented in this chapter supports the interpretation that linguistically down-regulated action sequences, mandatory in-game tasks, and multimodal presentation of vocabulary in LIS may be more effective at eliciting immediate gains in TL vocabulary acquisition and long-term retention after corrective feedback relative to the control intervention. Furthermore, comprehension of the narrative content was significantly correlated to vocabulary test scores. This positive correlation adds weight to the interpretation that effort toward comprehension while playing, corrective gameplay mechanics conducive to eliciting comprehension, and higher levels of starting English proficiency may lead to better TL vocabulary acquisition and retention in the LIS gaming intervention. Of particular note, however, was that the significant gender differences on comprehension tests was not reflected in vocabulary test scores. The following chapter will analyze in further detail sample variation within the LG group including, but not limited to, gender.

Chapter 6 Results for perceptions of smartphone gaming

The following chapter will detail the results of the data analysis which directly address RQ3: How will the LIS gaming intervention correlate to students' perceptions of and attitudes towards English smartphone gaming? Specifically, the following sections will present analyses from the pre- and post-test surveys from the LG group to address the research question above. Furthermore, within-group factors such as gender, smartphone gaming experience, and gaming location will be used to provide a clearer interpretation regarding LG group sample variation.

6.1 Analysis of pre-game smartphone engagement

The following analysis of the students' relationships to smartphone gaming was based on data gathered in the pre- and post-game surveys. Students were asked to rate how often they engage with smartphone gaming on a 10-point Likert scale (1=not at all, 10=frequently) in the pre-game survey, the results of which are summarized in Figure 6.1. Within-group, two-sample, two-tailed t-tests were used to test for significant differences in Likert values between the gender groups, for which no significant differences were observed (p=0.32). Students were also asked to rate how often they plan to engage with smartphone gaming after the study program on the same scale in the post-game survey (see Figure 6.2). Within-group, two-sample, two-tailed t-tests revealed no significant differences between the gender groups (p=0.74). Although there was a slight drop in post-game smartphone engagement Likert values relative to pre-game values in both the aggregate and male samples, between-group, two-sample, two-tailed t-tests revealed no significant differences for the aggregate sample (p=0.86) or the male sample (p=0.58). In contrast to the drop in Likert values observed in the aggregate and male samples, the female sample revealed a slight post-game gain relative to pre-game values. However, no significant difference was observed in the between-group *t*-test for the female samples (p=0.69).

Figure 6.1 Pre-game smartphone engagement Likert responses

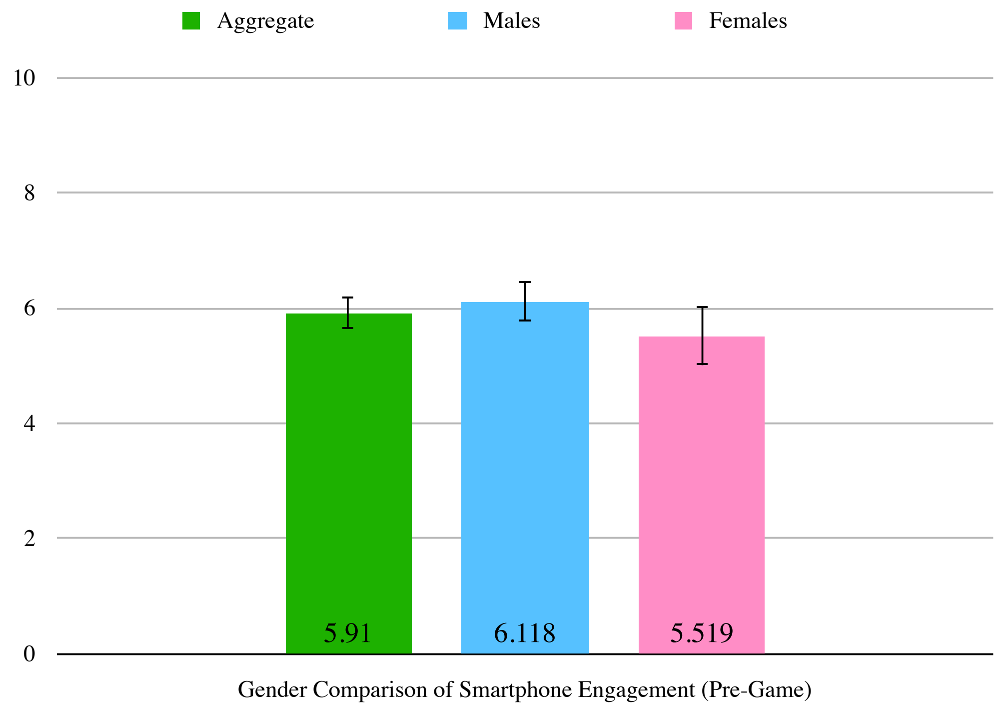
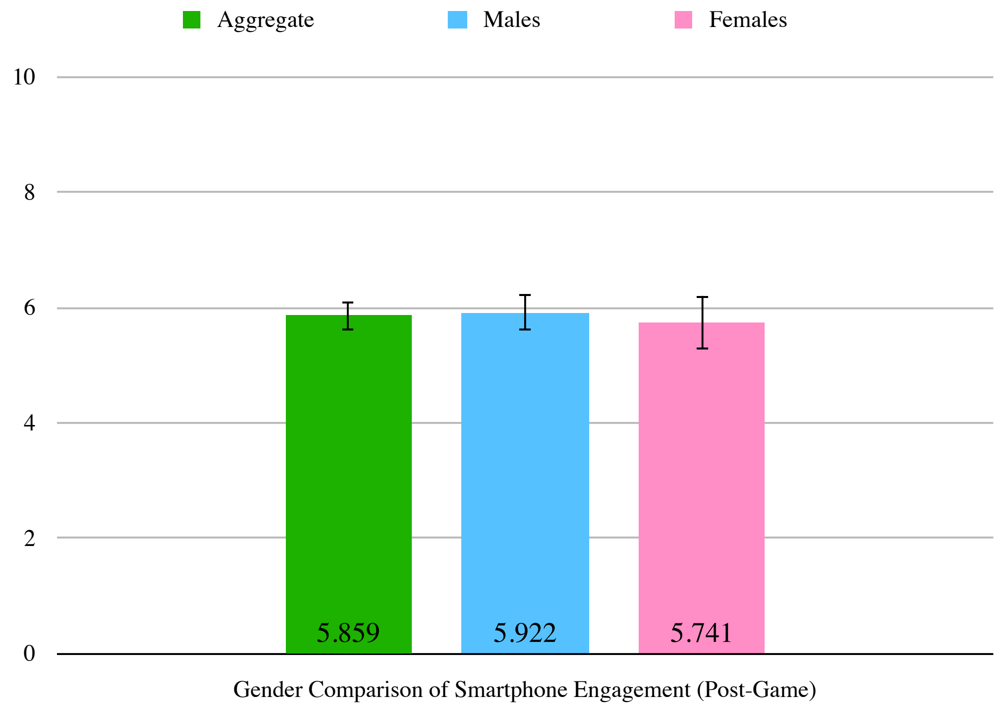


Figure 6.2 Post-game smartphone engagement Likert responses



Correlation analyses were also conducted for the pre-game smartphone engagement survey question to analyze potential correlations with other Likert scale survey questions as well as vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.1. Statistically significant correlations were revealed between the pre-game smartphone engagement values and post-game smartphone engagement values, post-game perception of smartphone gaming efficacy values, and perception of English growth values. Furthermore, significant differences in correlation coefficients between gender groups were observed in the perception of English growth values and post-game perception of smartphone gaming efficacy values. The following section will discuss the significant findings described above.

Table 6.1 Correlation analysis of pre-game smartphone engagement

	Aggregate (*n*=78)		Males (*n*=51)		Females (*n*=27)	
Categories	*r*-value	*p*-value	*r*-value	*p*-value	*r*-value	*p*-value
Vocabulary scores	-0.16	0.17	-0.2	0.16	-0.09	0.66
Comprehension scores	-0.23	0.15	-0.06	0.79	-0.37	0.17
Pre-game perception of smartphone gaming efficacy	0.09	0.45	-0.01	0.94	0.23	0.25
Content comprehension	0.20	0.08	0.21	0.13	0.13	0.51
Post-game smartphone engagement	0.40	0.00024	0.42	0.0023	0.38	0.05
Post-game perception of smartphone gaming efficacy	0.22	0.06	0.39	0.0049	-0.14	0.49
Content difficulty	-0.03	0.76	0.08	0.58	-0.20	0.32
Perception of English growth	0.07	0.54	0.30	0.033	-0.28	0.15
Categories with significant gender differences	Observed *z*-value		Fisher's *z* (Males)		Fisher's *z* (Females)	
Perception of English growth	2.40		0.31		-0.29	
Post-game perception of smartphone gaming efficacy	2.20		0.41		-0.14	

Moderate, significant correlations between pre-game smartphone engagement values and post-game smartphone engagement values were observed in the aggregate (r=0.4, p=0.00024) (see Figure 6.3), male (r=0.42, p=0.0023) (see Figure 6.4), and female samples (r=0.38, p=0.05) (see Figure 6.5). All observed significant correlations were positive, meaning that students, regardless of gender, who engaged with smartphone gaming before the study program were moderately more likely to plan to engage with smartphone gaming after completing the study program.

Figure 6.3 Pre-game smartphone engagement correlation analysis 1

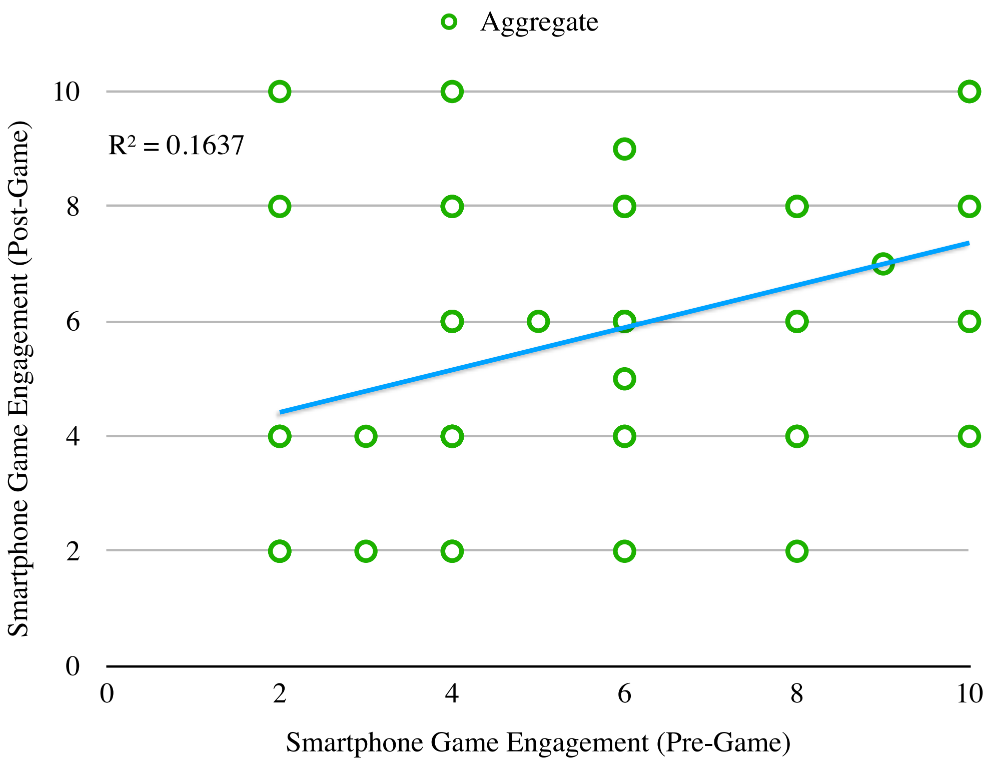


Figure 6.4 Pre-game smartphone engagement correlation analysis 2

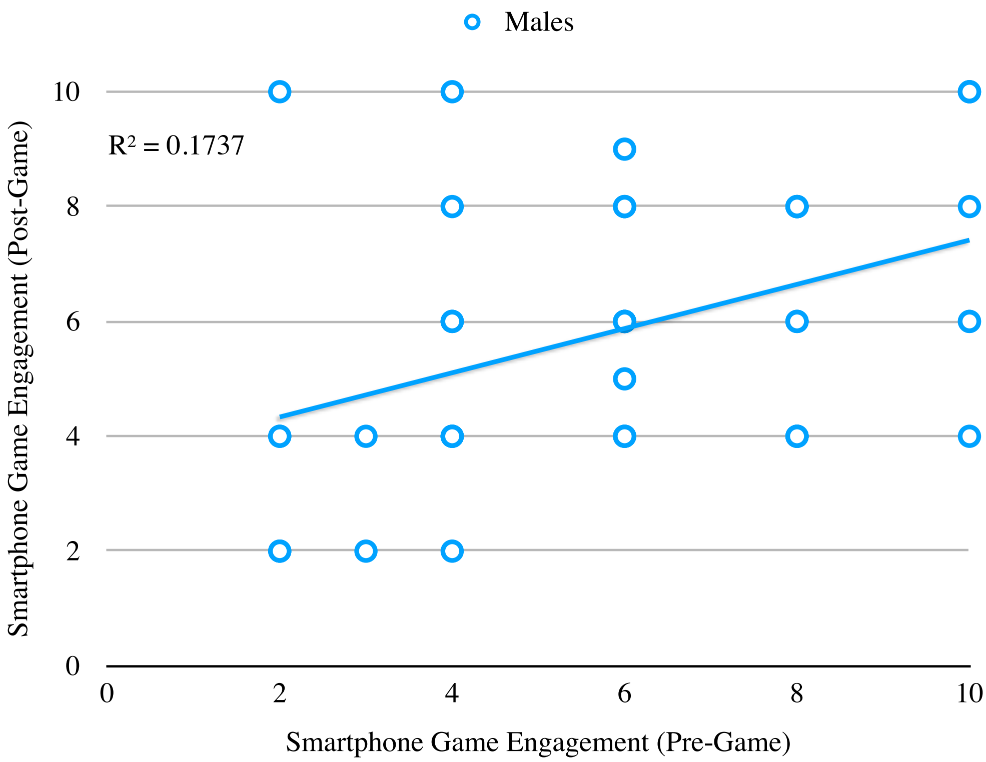
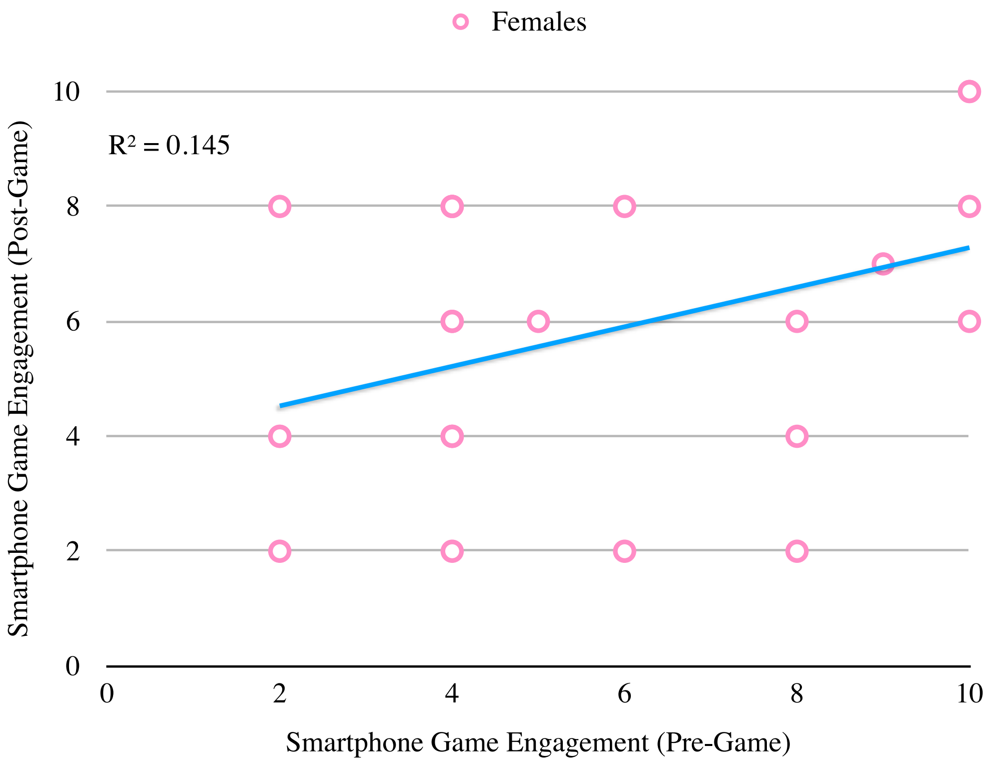


Figure 6.5 Pre-game smartphone engagement correlation analysis 3



A moderate positive correlation between pre-game smartphone engagement values and post-game perception of smartphone gaming efficacy was observed in the male sample (r=0.39, p=0.0049) (see Figure 6.6). The positive correlation provides limited evidence to support the argument that the males who engaged with smartphone gaming more frequently were more likely to perceive smartphone gaming as an effective language learning tool after participating in the present study. This correlation was significantly different in the female sample. A non-significant, negligible negative correlation (r=-0.14, p=0.49) between the above-stated survey values was observed in the female sample (see Figure 6.7). Although this correlation in the female sample is non-significant and negligible, the differences in r-values between the gender groups was observed to be significant (z=2.20, p=0.028). With this significant finding in mind, the argument may be made that the pre-study smartphone gaming habits of females were less likely than those of males to affect how efficacious they perceived smartphone gaming to be as a language learning tool after participating in the present study.

Figure 6.6 Pre-game smartphone engagement correlation analysis 4

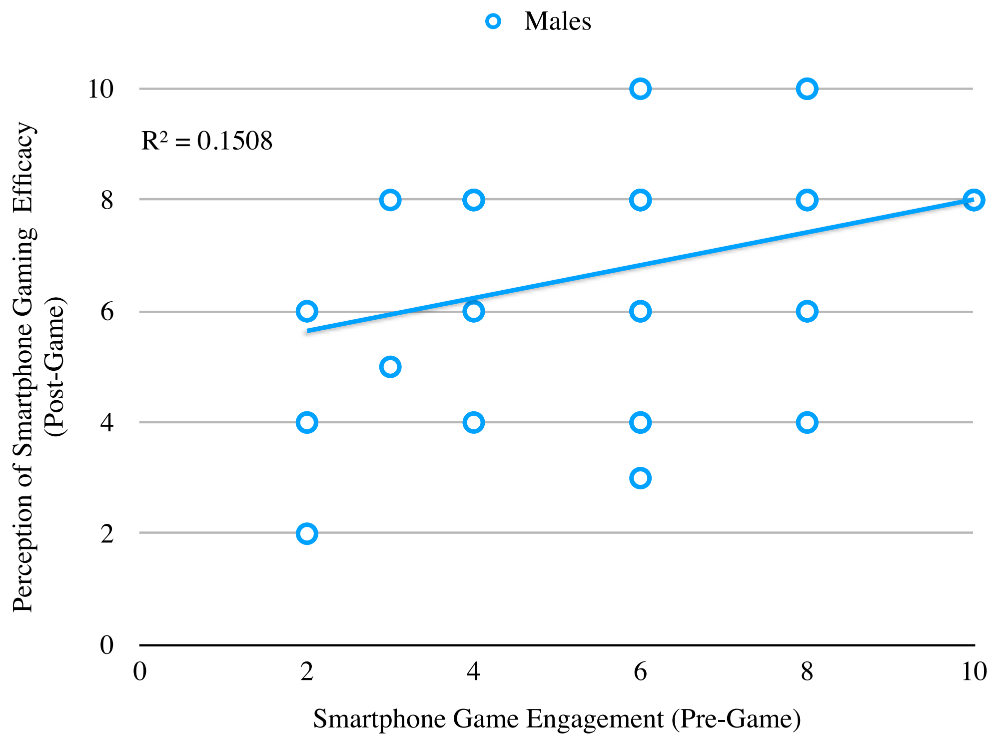
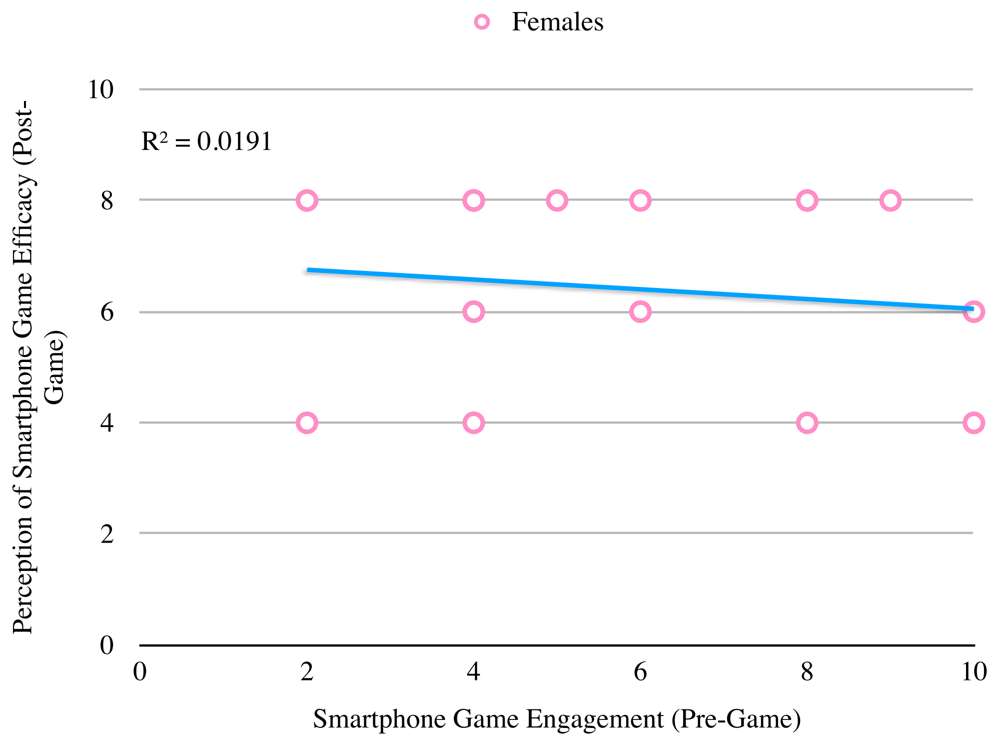


Figure 6.7 Pre-game smartphone engagement correlation analysis 5



Finally, a moderate positive correlation was observed between pre-game smartphone engagement values and perception of English growth values in the male sample (r=0.3, p=0.033) (Figure 6.8). This positive correlation provides limited evidence to make the argument that the males who engaged with smartphone gaming more frequently pre-study were more likely to perceive that their English language skill grew as a result of participating in this study. This correlation was significantly different in the female sample. A non-significant, negligible negative correlation (r=-0.28, p=0.15) between the above-stated survey values was observed in the female sample (see Figure 6.9). Although this correlation in the female sample is non-significant and negligible, the differences in r-values between the gender groups was observed to be significant (z=2.40, p=0.016). With this significant finding in mind, the argument may be made that the smartphone gaming habits of females were less likely than those of males to affect how much they perceived their English skills to grow as a result of participating in the present study.

Figure 6.8 Pre-game smartphone engagement correlation analysis 6

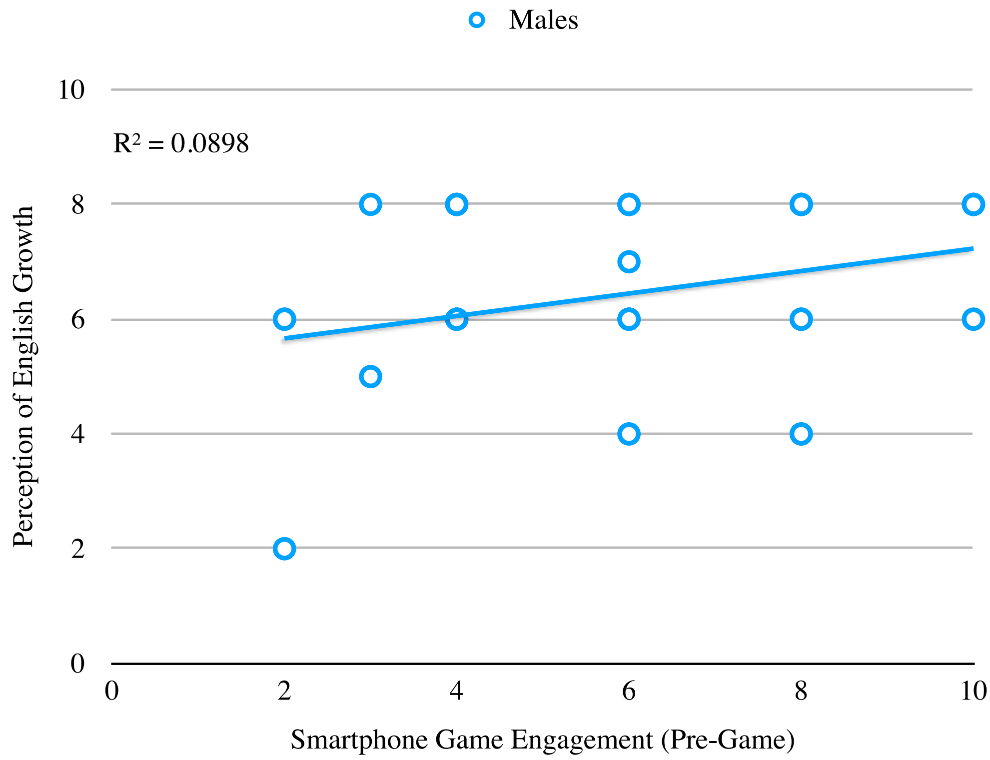
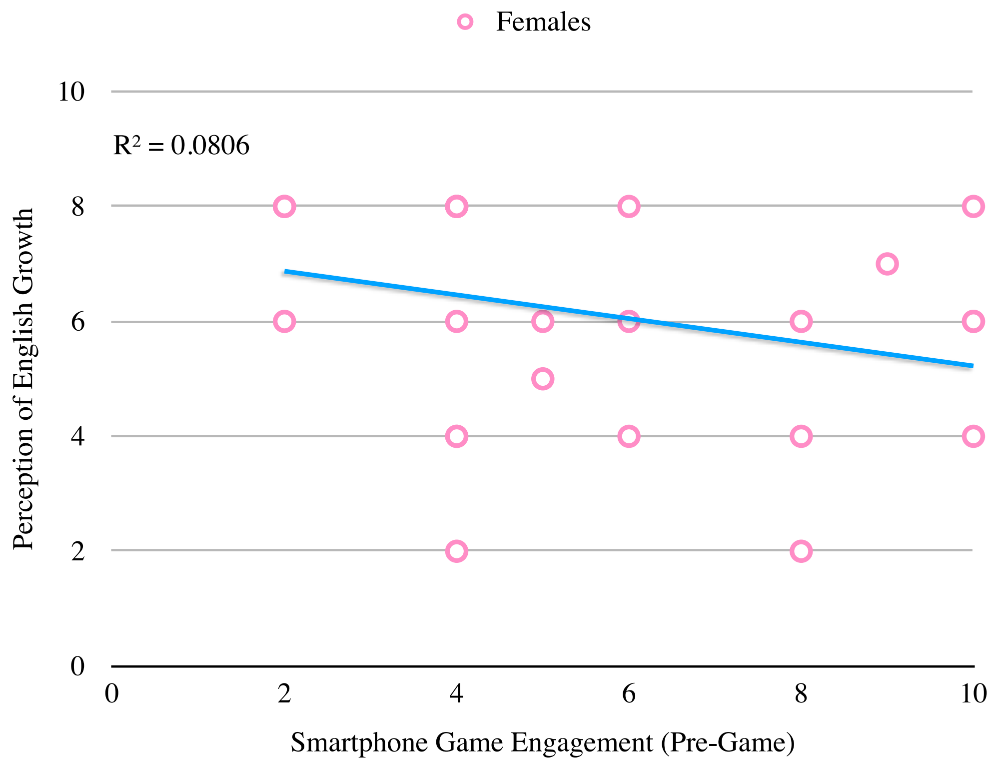


Figure 6.9 Pre-game smartphone engagement correlation analysis 7



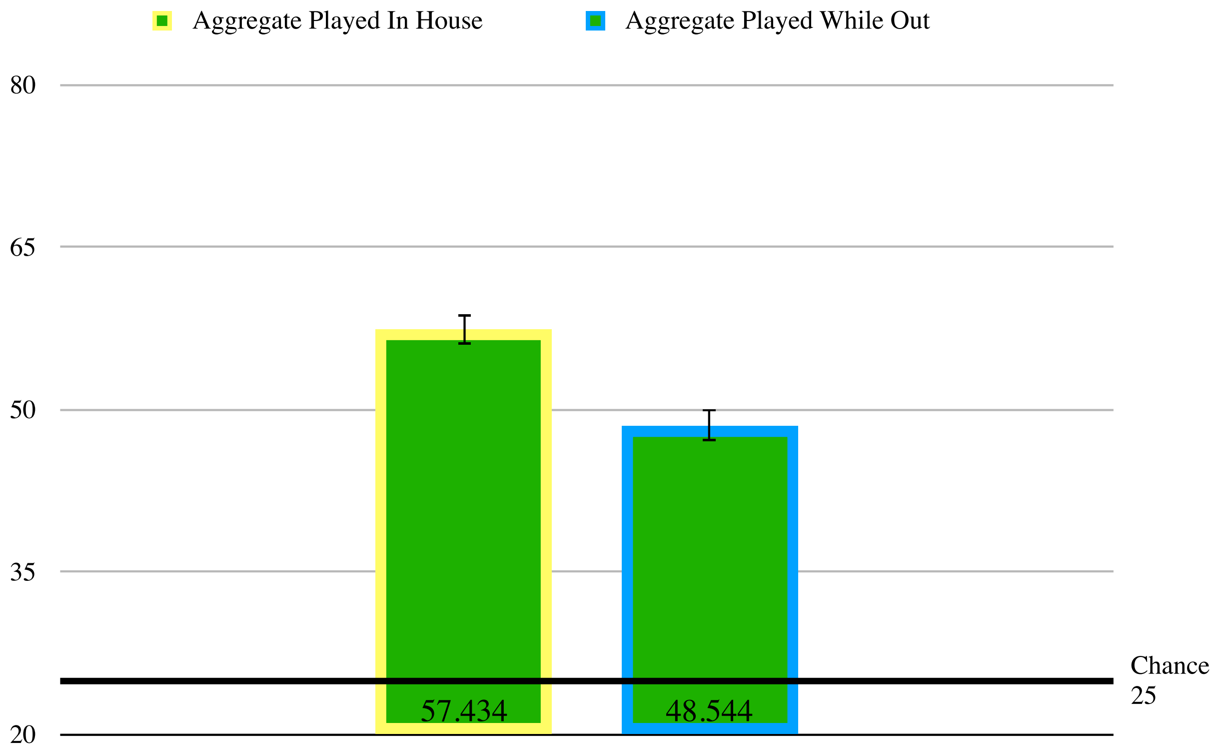
6.2 Analysis of smartphone gaming location

Due to the effect of the Covid-19 pandemic, the gaming protocol included gaming sessions both in and out of the classroom. In order to determine whether the location at which the participants in this study played LIS affected their vocabulary test scores, the post-game survey collected data on where participants played LIS during the study. Three categories of location were identified in the student responses: at home, in the classroom, and at an outside location (e.g., at a café or in-transit). To analyze whether these gaming locations affected the students' vocabulary test scores, a one-way ANOVA analysis was performed to detect any statistically significant differences between the above-stated location groups. Furthermore, between-group, two-sample *t*-tests were run between each location category, and within-group, two-sample *t*-tests were run on the gender groups within each location category. The results are summarized in Table 6.2. The one-way ANOVA analysis produced an *f*-value of 3.52 indicating statistically significant variation in the vocabulary test results from the three gaming location categories. The between-group *t*-tests provide further detail as to the source of the significant variation. There was a moderately significant difference (p=0.066) between the aggregate sample in the at-home and outside/in-transit locations (see Figure 6.10). This difference was more pronounced in the female sample (p=0.086) than in the male sample (p=0.39), providing limited evidence that the female participants in this study were more likely to benefit from engaging with the gaming protocol at home. No other significant results were observed in the analyses described above.

Table 6.2 Analysis of gaming locations (post-game)

	Aggregate (n=78)	Males (n=51)	Females (n=27)
Vocabulary test averages for gaming and testing locations			
Home	n=52 57.43	n=32 55.92	n=20 59.85
Classroom	n=19 54	n=14 54.13	n=5 53.64
Outside/in transit	n=8 48.54	n=5 51.01	n=3 44.44
ANOVA analysis			f = 3.52
Between-group t-tests			
Home to classroom	0.30	0.61	0.4
Classroom to outside/in transit	0.29	0.61	0.42
Home to outside/in transit	0.066	0.39	0.086
Within-group two-sample gender t-tests			
Home			0.27
Classroom			0.94
Outside/in transit			0.51

Figure 6.10 Gaming location analysis



One potential source of the slightly higher vocabulary test averages in the at-home location category relative to the in-class and outside/in-transit categories may be found in the students' responses to the pre-game survey gaming location question (see Figure 6.11 and Figure 6.12). A significant percentage (55%) of responses to the pre-game location question indicated that students typically play smartphone games at home as opposed to the other categories. The general familiarity with the gaming location may have served to bolster in-game comprehension and thus vocabulary acquisition and retention while participating in the present study. Other possible interpretations include the arguably higher levels of distraction when playing smartphone games while out or in transit. However, this explanation is purely speculative and there was no quantitative or qualitative data collected in the present study to measure how distractions in different gaming locations may or may not affect the acquisition and retention of in-game vocabulary. Of further note is the disproportionately large percentage of males who play smartphone games while out (85.71%) and at work (80%)

relative to females (14.29% and 20% respectively) relative to the sample size (males=65.39%, females=34.62%).

Figure 6.11 Pre-game smartphone gaming locations

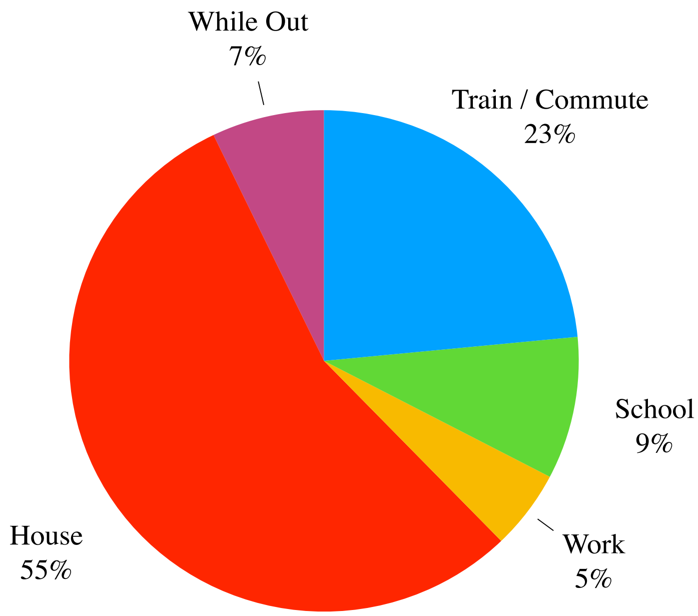
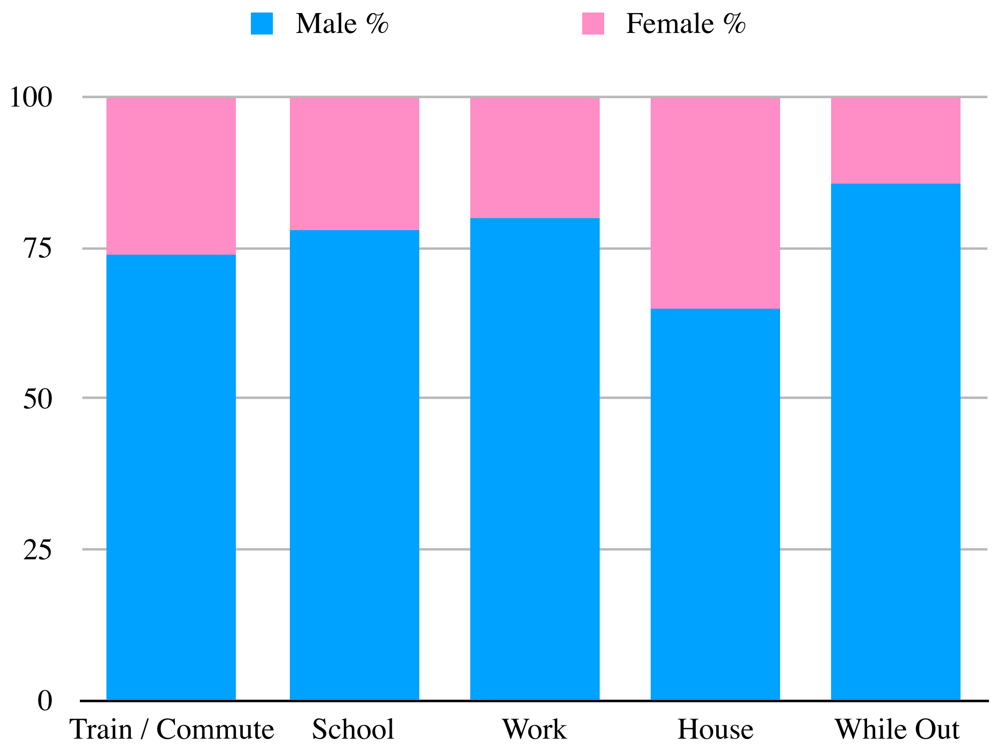


Figure 6.12 Pre-game smartphone gaming locations gender breakdown



6.3 Analysis of prior English smartphone gaming experience

In order to determine whether prior experience playing smartphone games in English affected vocabulary test scores, the pre-game survey collected data on whether the students had played a smartphone game in English (and what game) before participating in the present study program. Three categories of responses were identified: prior experience playing smartphone games in English other than LIS, prior experience playing LIS in English (primarily from the multiple play-through sample), and no prior experience playing smartphone games in English. To analyze whether prior English smartphone gaming experience affected the students' vocabulary test scores, a one-way ANOVA analysis was performed to detect any statistically significant differences between the above-stated English gaming experience groups. Furthermore, between-group, two-sample *t*-tests were run between each English gaming experience category, and within-group, two-sample *t*-tests were run on the gender groups within each experience category. The results are summarized in Table 6.3. The one-way ANOVA analysis produced an *f*-value of 5.88 indicating statistically significant variation in the vocabulary test results from the three gaming experience categories. The between-group *t*-tests provide further detail as to the source of the significant variation. There were moderately significant differences between the aggregate samples of prior English gaming experience relative to no English gaming experience (p=0.065) and prior LIS English gaming experience relative to no English gaming experience (p=0.064) (see Figure 6.13). Furthermore, there were also statistically significant differences between the male samples of the categories described above (p-values of 0.024 and 0.043 respectively) (see Figure 6.14). No other significant results were observed in the analyses described above. The significant results in the male sample for the prior English smartphone gaming experience category add further weight to the correlation analysis data suggesting that males who engage with English smartphone gaming more frequently were more likely to perform better in the testing protocol used in the present study.

Table 6.3 Analysis of prior English smartphone gaming experience

	Aggregate (*n*=78)	Males (*n*=51)	Females (*n*=27)
Vocabulary test averages for gaming experience categories			
English gaming experience	*n*=13 60.59	*n*=8 62.1	*n*=5 58.16
LIS experience in English	*n*=8 62.1	*n*=6 62.02	*n*=2 62.33
No English gaming experience	*n*=57 53.49	*n*=37 52.41	*n*=20 55.49
ANOVA analysis			*f* = 5.88
Between-group *t*-tests			
English exp. to LIS	0.75	0.99	0.5
LIS exp. to no exp.	0.064	0.043	0.07
No exp. to English	0.065	0.024	0.73
Within-group two-sample gender *t*-tests			
English exp.			0.57
LIS exp.			0.94
No English exp.			0.38

Figure 6.13 Prior English smartphone gaming experience vocabulary test averages (aggregate sample)

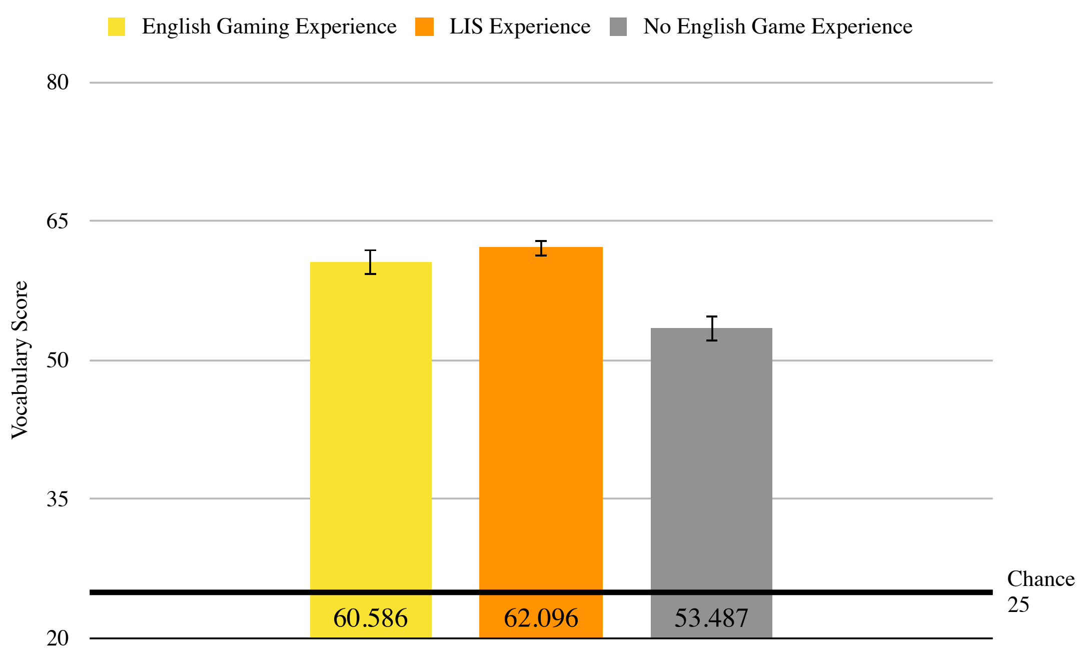
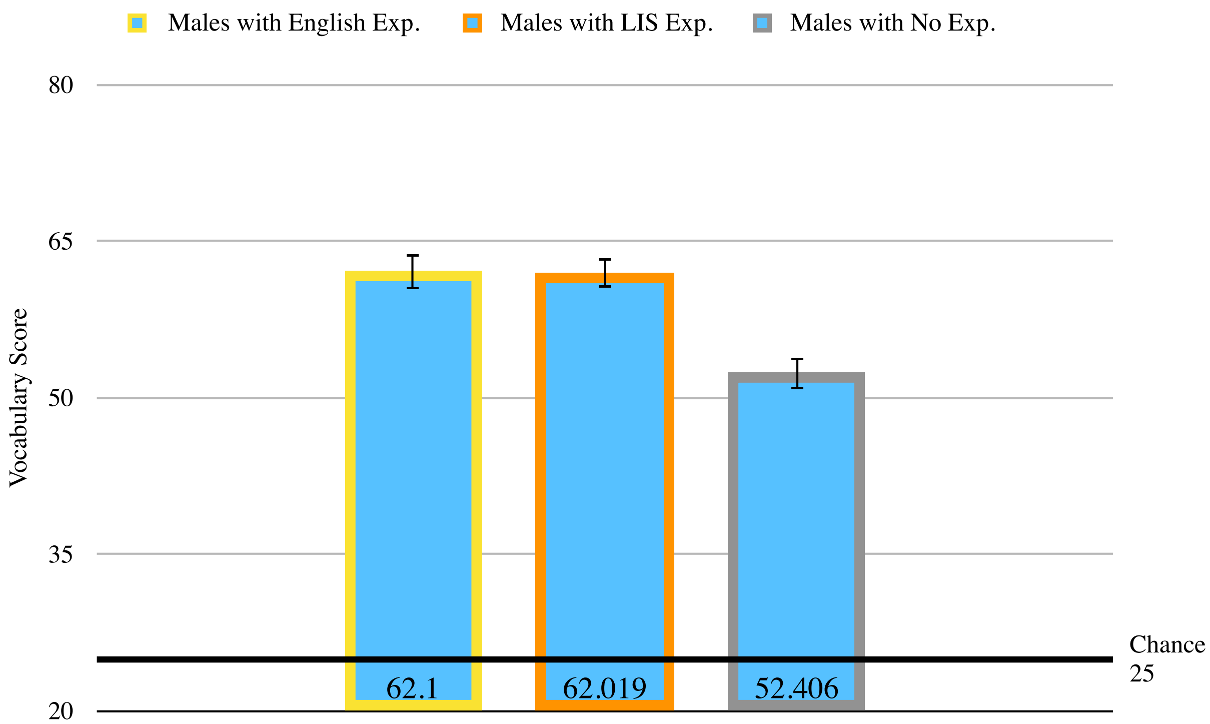


Figure 6.14 Prior English smartphone gaming experience vocabulary test averages (male sample)



6.4 Analysis of language used with prior LIS experience

Due to the possibility of having participants in the study who had prior LIS gaming experience as well as repeat participants in the study program, the pre-game survey collected data on whether the students had played LIS before and if so, in what language. Three categories of responses were identified for analysis: prior experience playing LIS in English, prior experience playing LIS in Japanese or another language (e.g., Chinese), and no prior experience playing LIS. To analyze whether prior experience playing LIS affected the students' vocabulary test scores, a one-way ANOVA analysis was performed to detect any statistically significant differences between the above-stated English gaming experience groups. Furthermore, between-group, two-sample *t*-tests were run between each LIS gaming experience category, and within-group, two-sample *t*-tests were run on the gender groups within each LIS experience category. The results are summarized in Table 6.4. The one-way

ANOVA analysis produced an *f*-value of 7.35 indicating highly statistically significant variation in the vocabulary test results from the three LIS gaming experience categories. The between-group *t*-tests provide further detail as to the source of the significant variation.

Table 6.4 Analysis of language used with prior LIS experience

	Aggregate (n=78)	Males (n=51)	Females (n=27)
Vocabulary test averages for language used with prior LIS experience			
English	n=12 63.26	n=10 63.44	n=2 62.33
Japanese/other	n=3 60.95	n=3 60.95	n=0
No LIS experience	n=63 53.83	n=38 52.39	n=25 56.02
ANOVA analysis			f = 7.35
Between-group t-tests			
English to Japanese/other	0.68	0.69	
Japanese/Other to no LIS exp.	0.34	0.19	
No LIS exp. to English	0.015	0.0047	0.045
Within-Group Two-Sample Gender t-tests			
English			0.7
Japanese/other			
No LIS experience			0.27

There was a significant difference between the aggregate samples of prior LIS gaming experience in English relative to no LIS gaming experience (p=0.015) (see Figure 6.15). Furthermore, there was also a highly statistically significant difference between the male samples of the category described above (p=0.0047) (see Figure 6.16), although a significant difference (p=0.045) was also observed in the female sample in the same category to a lesser degree. No other significant results were observed. The highly significant results in the male sample for the prior LIS gaming experience in English add further weight to the correlation analysis data suggesting that males who engage with English smartphone gaming more frequently were more likely to perform better in the testing protocol used in the present study.

Figure 6.15 Prior LIS gaming experience vocabulary test averages (aggregate sample)

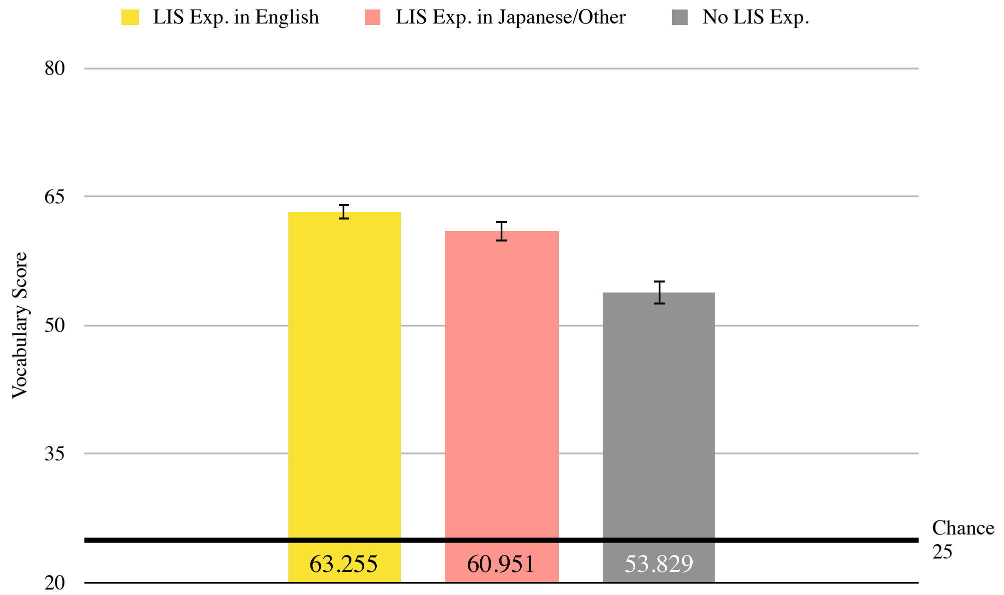
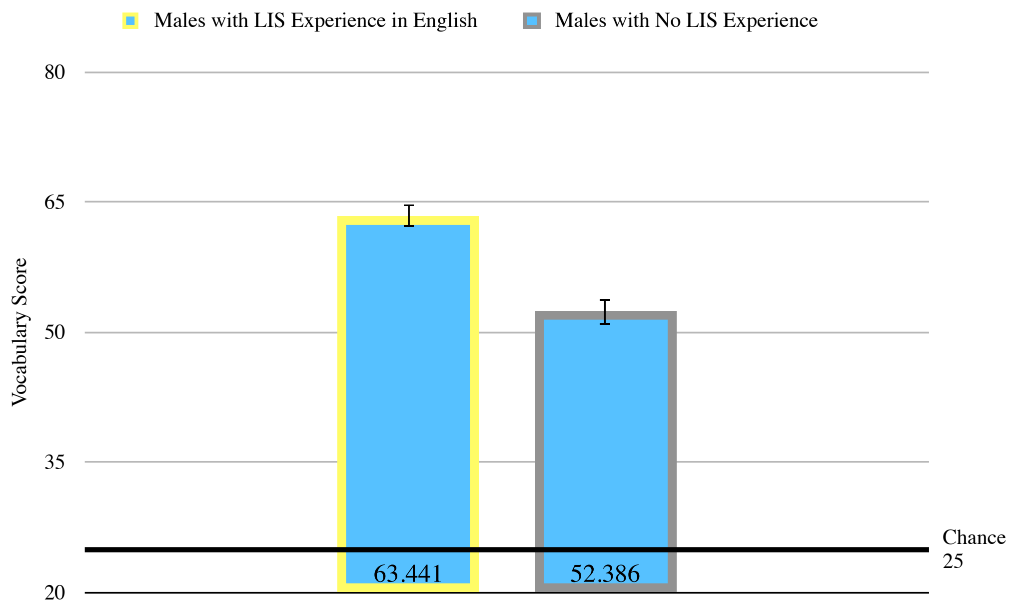


Figure 6.16 Prior LIS gaming experience vocabulary test averages (male sample)



6.5 Analysis of pre-game perception of smartphone gaming efficacy

The following analysis of the students' perceptions of smartphone gaming efficacy for language learning purposes was based on data gathered in the pre- and post-game surveys. Students were asked to rate how efficacious they perceived smartphone games to be for the purpose of learning English on a 10-point Likert scale (1=not at all, 10=very efficacious), the results of which are summarized in Figure 6.17 (pre-game) and Figure 6.18 (post-game). Although there were slight drops in the post-game perception of smartphone gaming efficacy Likert values relative to pre-game values in all samples, two-sample, two-tailed *t*-tests revealed no significant differences.

Figure 6.17 Pre-game perception of smartphone gaming efficacy Likert responses

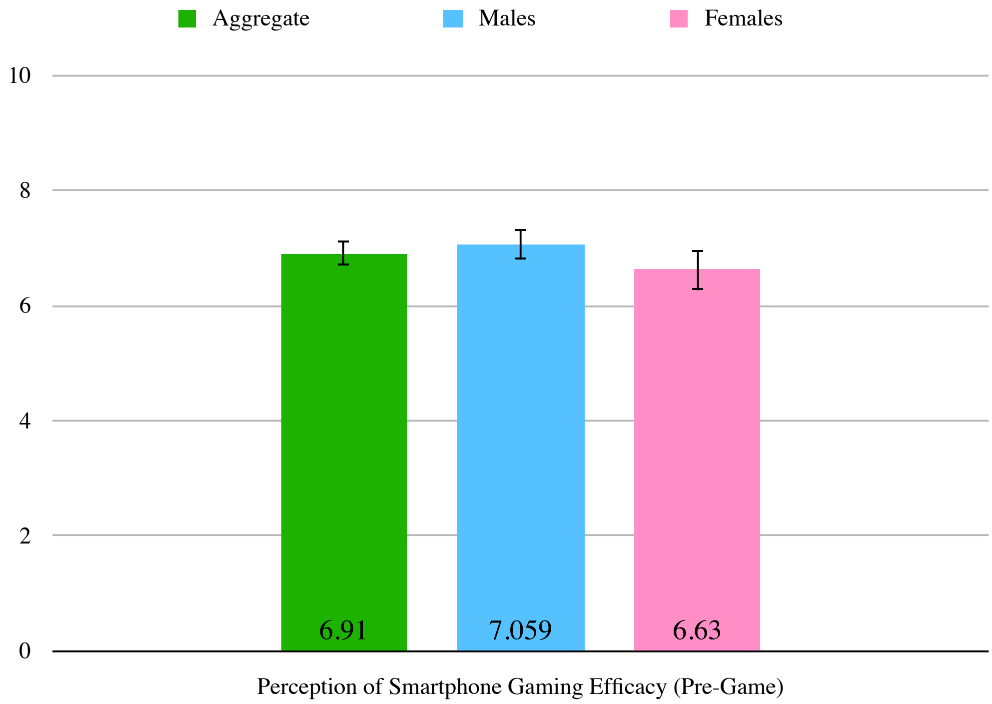
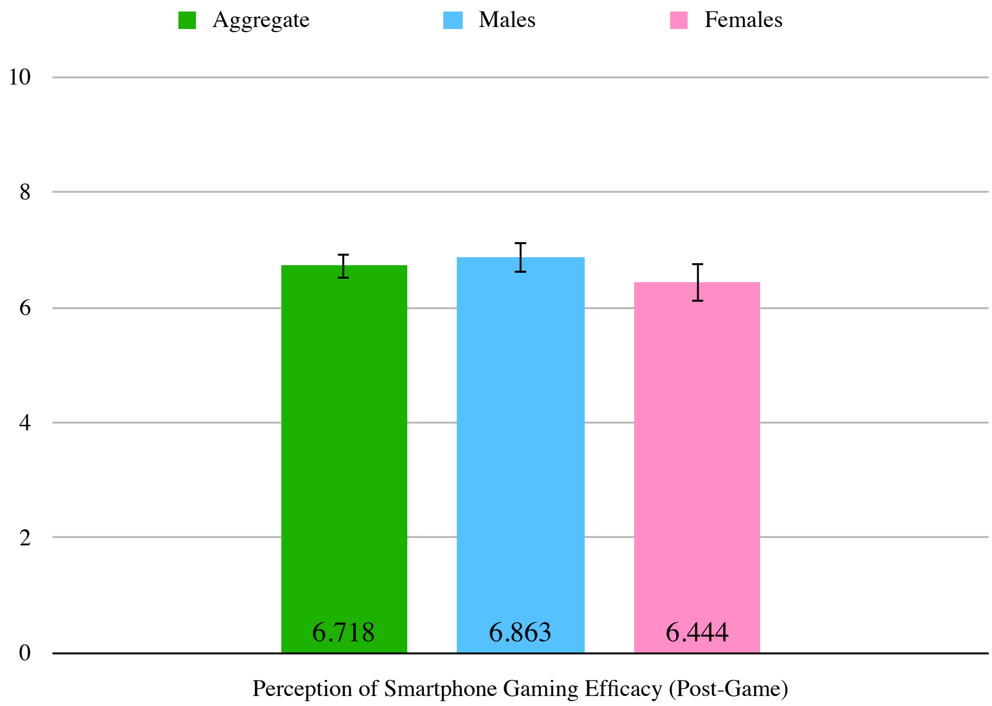


Figure 6.18 Post-game perception of smartphone gaming efficacy Likert responses



Correlation analyses were also conducted for the pre-game perception of smartphone gaming efficacy survey question to analyze potential correlations with other Likert scale survey questions as well as with vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.5. Statistically significant correlations were revealed between the pre-game perception of smartphone gaming efficacy values and vocabulary scores, comprehension scores, post-game perception of smartphone gaming efficacy values, and perception of English growth values. No significant differences in correlation coefficients between gender groups were observed in the correlation analyses described above. The following section will detail the significant findings.

Table 6.5 Correlation analysis of the perception of smartphone gaming efficacy (pre-game)

	Aggregate (n=78)		Males (n=51)		Females (n=27)	
Categories	r-value	p-value	r-value	p-value	r-value	p-value
Vocabulary scores	-0.24	0.032	-0.33	0.019	-0.11	0.59
Comprehension scores	-0.44	0.0035	-0.40	0.040	-0.67	0.0063
Content comprehension	-0.0017	0.99	0.037	0.80	-0.13	0.52
Post-game smartphone engagement	0.18	0.11	0.13	0.38	0.28	0.15
Post-game perception of smartphone gaming efficacy	0.32	0.0042	0.42	0.0022	0.08	0.69
Content difficulty	0.11	0.35	0.10	0.47	0.12	0.54
Perception of English growth	0.25	0.028	0.28	0.046	0.18	0.37

Moderately weak correlations were observed between the pre-game perception of smartphone gaming efficacy values and vocabulary test scores in both the aggregate sample (r=-0.24, p=0.032) (see Figure 6.19) and male sample (r=-0.33, p=0.019) (see Figure 6.20). Both observed correlations are negative, potentially indicating that students who perceived smartphone gaming to be an efficacious language learning tool prior to participating in the study program were slightly less likely to perform well on subsequent vocabulary tests. This negative correlation was also observed in the female sample, albeit less pronounced and non-significant (r=-0.11, p=0.59). The male sample in this study may therefore be more susceptible to this particular correlation.

Figure 6.19 Pre-game perception of smartphone gaming efficacy analysis 1

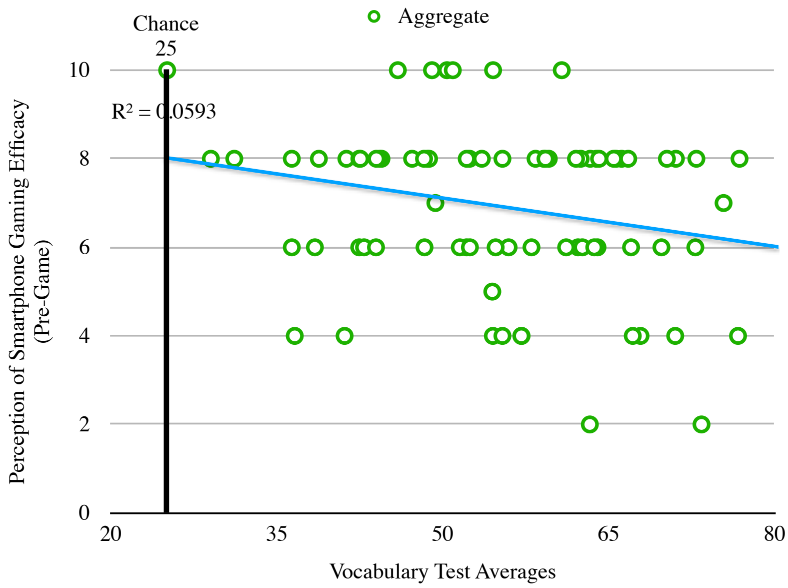
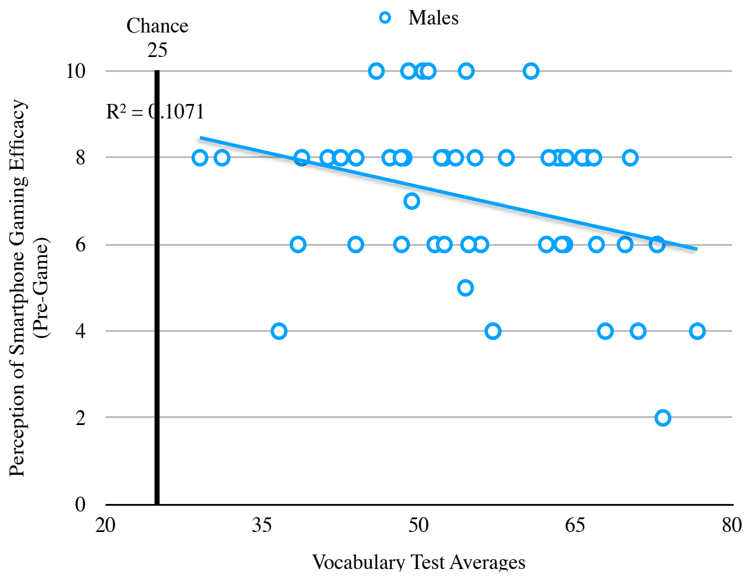


Figure 6.20 Pre-game perception of smartphone gaming efficacy analysis 2



Negative correlations between the pre-game perception of smartphone gaming efficacy values and comprehension test scores were also observed in all three sample groups. Moderate correlations were observed in the aggregate sample (r=-0.44, p=0.0035) (see Figure 6.21) and male sample (r=-0.40, p=0.040) (see Figure 6.22), and a strong correlation was observed in the female sample (r=-0.67, p=0.0063) (see Figure 6.23). Just as with the correlation to the vocabulary test scores, students who perceived smartphone gaming as an efficacious tool pre-study were moderately less likely to perform well on subsequent comprehension tests. In contrast to the vocabulary test correlation analysis, the correlation between perceived smartphone gaming efficacy for language learning purposes (pre-study) and comprehension test scores was more pronounced in females.

Figure 6.21 Pre-game perception of smartphone gaming efficacy analysis 3

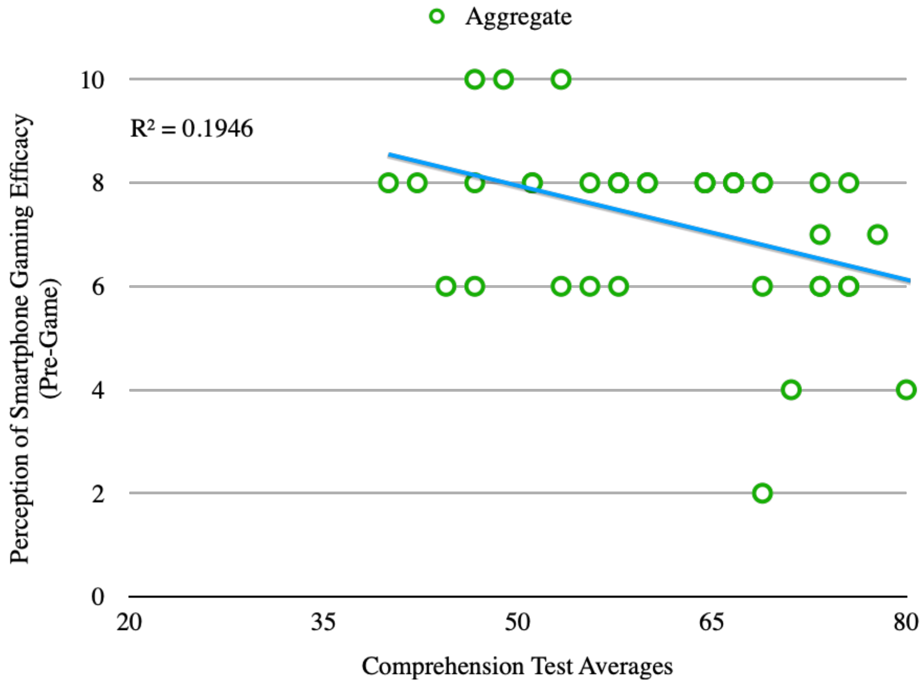


Figure 6.22 Pre-game perception of smartphone gaming efficacy analysis 4

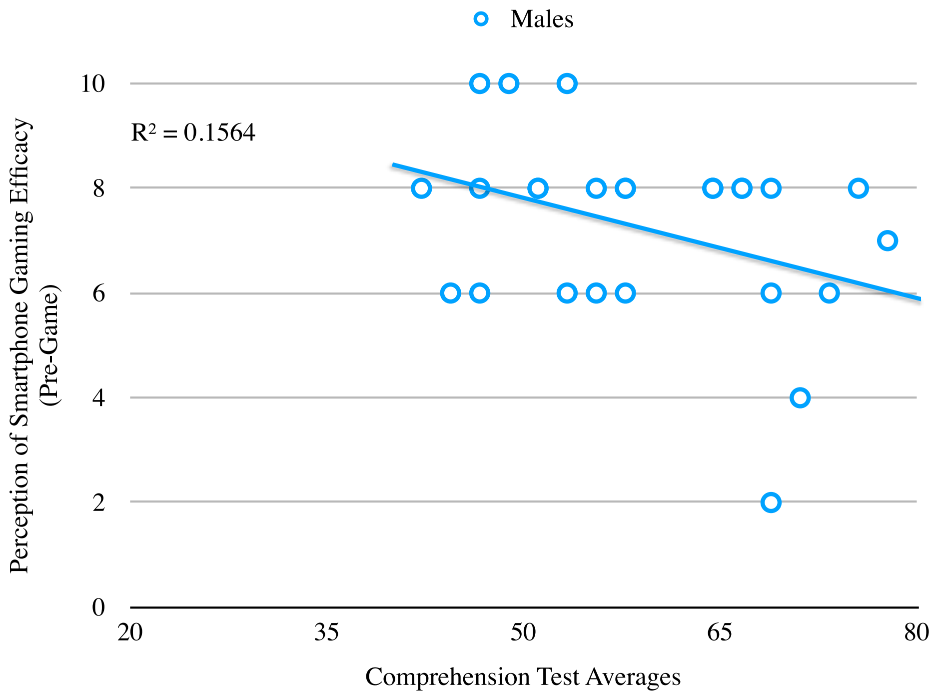
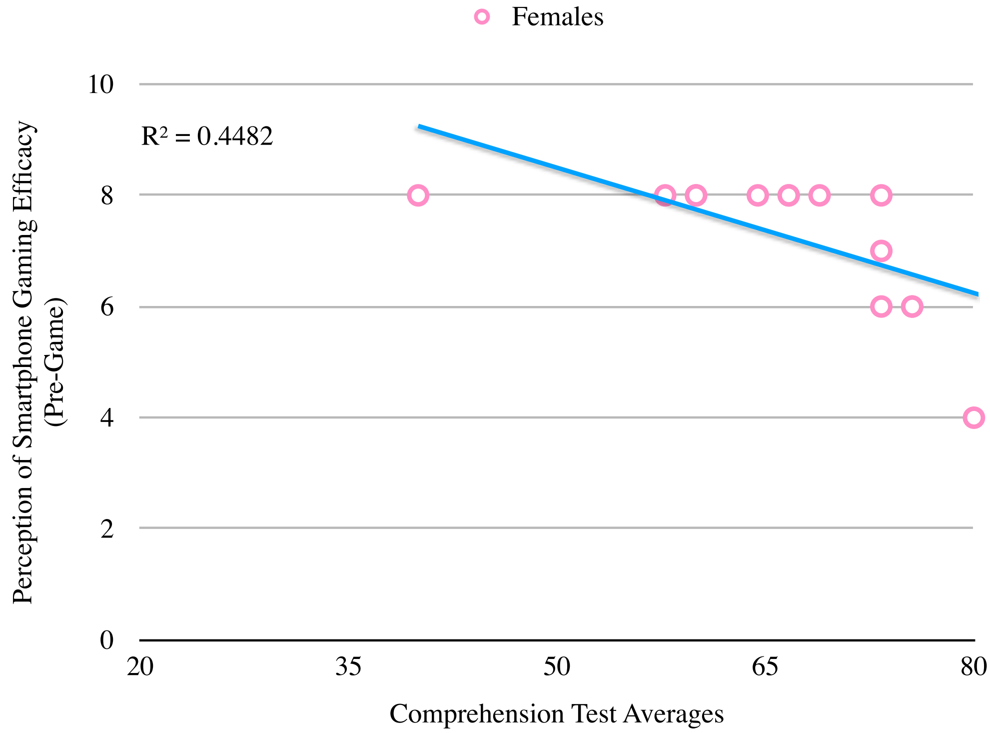


Figure 6.23 Pre-game perception of smartphone gaming efficacy analysis 5



Positive correlations were observed between the pre-game perceptions of smartphone gaming efficacy values and the post-game perceptions of smartphone efficacy values in the aggregate and male samples. The aggregate sample displayed an *r*-value of 0.32 with a *p*-value of 0.0042 (see Figure 6.24) and the male sample displayed an *r*-value of 0.42 with a *p*-value of 0.0022 (see Figure 6.25). The moderate positive correlation is contrasted by a non-significant, negligible correlation in the female sample (r=0.08, p=0.69). The differences between the two *r*-values was not observed to be statistically significant. As both gender groups' perceptions of smartphone gaming efficacy for language learning purposes were approximately the same both pre- and post-study, the significant, moderate correlation in the male sample and lack of significant correlation in the female sample is noteworthy. A plausible explanation for these results is that a moderate number of males in the dataset who were more likely to perceive smartphone gaming as an efficacious language learning tool before participating in the study program were also more likely to perceive them to be efficacious language learning tools after participating in the study program. For the female sample, however, a number of participants likely changed their perceptions of smartphone gaming efficacy from pre- to post-game, which would preserve the total Likert averages from pre- to post-game and result in a negligible correlation between the pre- and post-game smartphone gaming efficacy values. This indicates that a certain number of females who perceived smartphone gaming as an efficacious language learning tool before participating in the study were less likely to perceive them as efficacious language learning tools after participating in the study (and vice versa).

Figure 6.24 Pre-game perception of smartphone gaming efficacy analysis 6

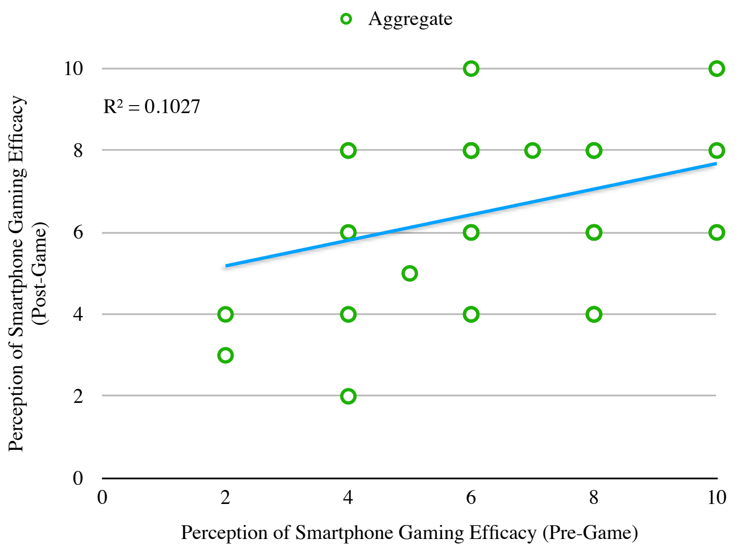
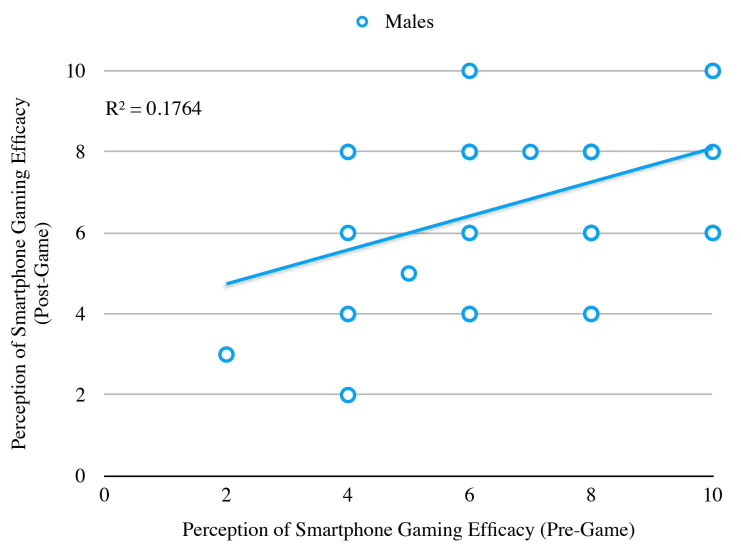


Figure 6.25 Pre-game perception of smartphone gaming efficacy analysis 7



Finally, there were weak positive correlations observed between the perception of smartphone gaming efficacy values and perception of English growth values in both aggregate and male samples. The aggregate sample displayed an *r*-value of 0.25 with a *p*-value of 0.028 (see Figure 6.26) and the male sample displayed an *r*-value of 0.28 with a *p*-value of 0.046 (see Figure 6.27), indicating that males who perceived smartphone gaming to be an efficacious language learning tool before participating in this study were more likely to perceive that their English proficiency grew as a result of playing LIS. This positive correlation was also observed in the female sample; however, the correlation was determined to be negligible and non-significant (r=0.18, p=0.37), which may provide limited evidence that the correlation described above was slightly more pronounced in the male population. This data interpretation, however, is highly speculative as the differences between *r*-values in both gender groups was observed to be non-significant.

Figure 6.26 Pre-game perception of smartphone gaming efficacy analysis 8

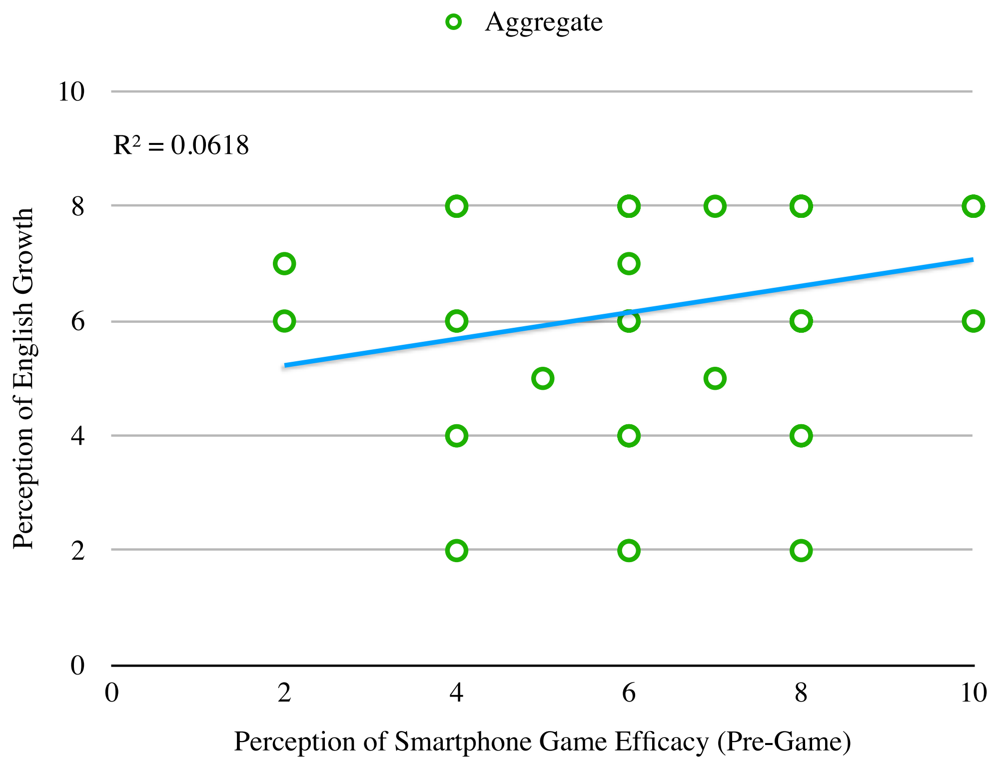
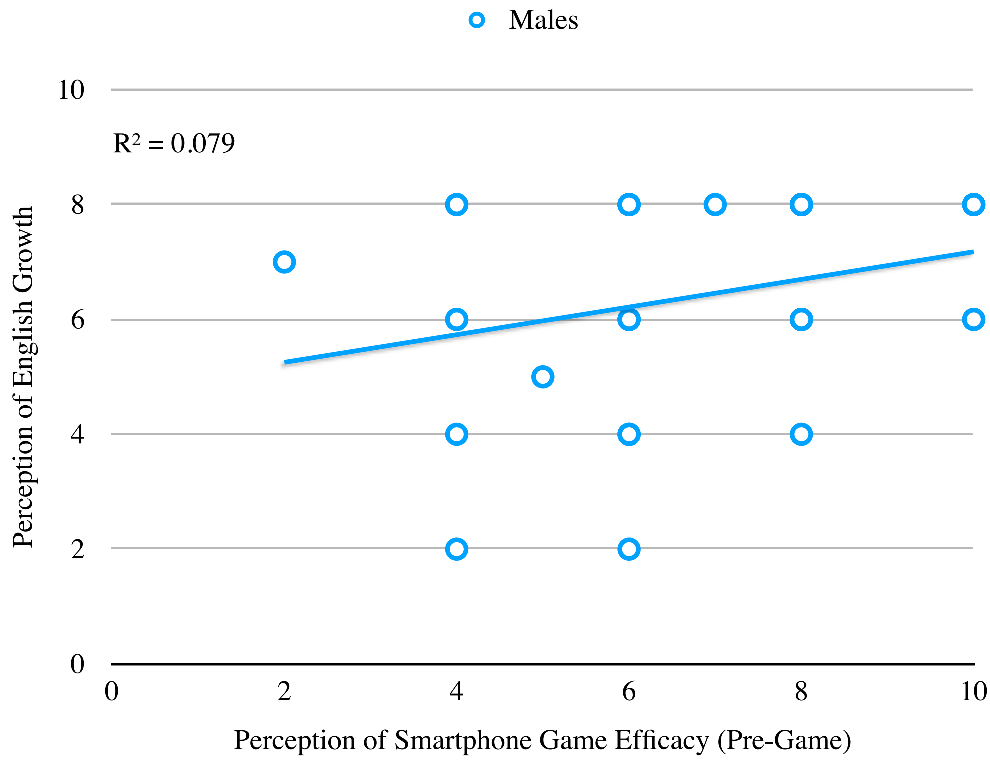


Figure 6.27 Pre-game perception of smartphone gaming efficacy analysis 9



6.6 Analysis of the perception of linguistic improvement form smartphone gaming

In order to measure how the students perceived their English proficiency to be improved with the LIS gaming intervention used in the present study, a pre-game survey question asked students to list what linguistic skills they thought would be improved by playing smartphone games in English. This pre-game survey item was followed up with an identical post-game survey question. There were four categories of student responses recorded from the pre-game survey: 1) vocabulary, 2) listening, 3) speaking, and 4) reading. The percentage breakdown for each response category is summarized in Figure 6.28 and the gender percentage breakdown for each response category is summarized in Figure 6.29. Overall, the largest response category was vocabulary (40%), followed by listening (29.66%), reading (18.62%), and speaking (11.72%). There were a total of 51 males in the sample (65.39%) and 27 females (34.62%). The gender breakdown for each response category revealed an approximately equal spread of responses for each category relative to the gender percentage for the dataset.

Figure 6.28 Perception of linguistic improvement from smartphone gaming category breakdown (pre-game)

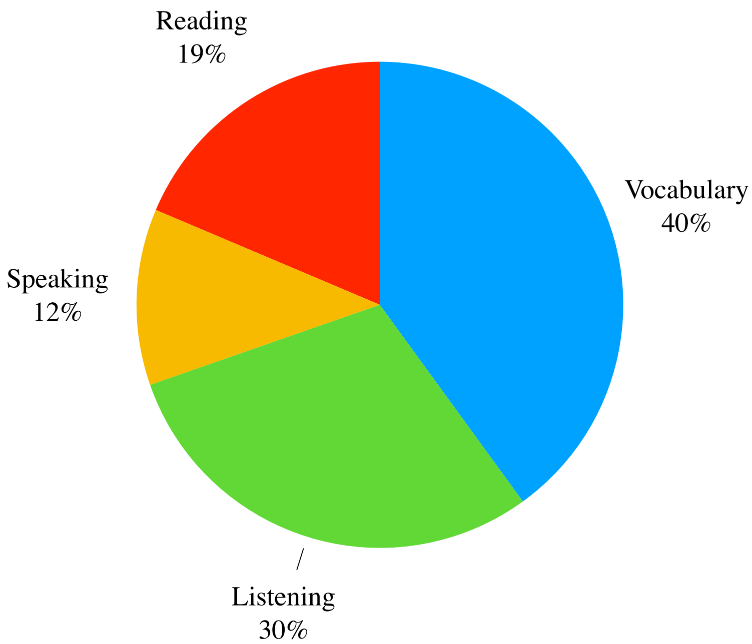
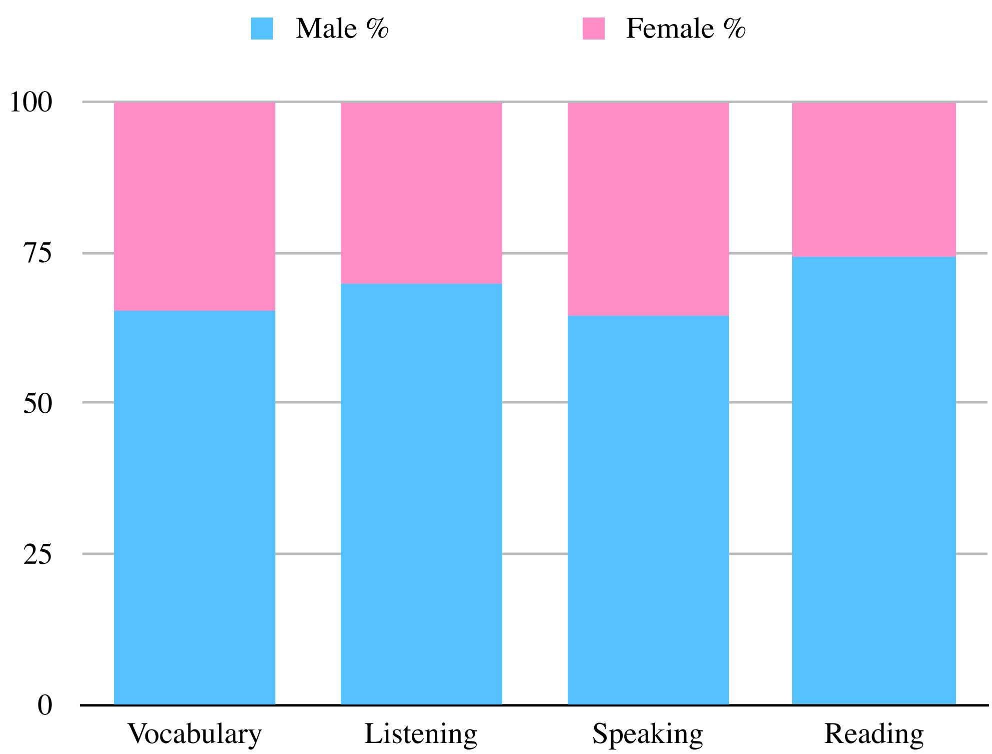


Figure 6.29 Perception of linguistic improvement from smartphone gaming gender breakdown (pre-game)



The post-test survey data for the perception of linguistic improvement responses are summarized in Figure 6.30, and the gender breakdown is summarized in Figure 6.31. The percentage changes in category responses were analyzed using paired, two-tailed *t*-tests. A significant increase in the percentage of students who identified listening as a skill they thought would be improved by playing smartphone games in English from pre-game (29.66%) to post-game (37.11%) was observed (p=0.004). Although there was a nine-point drop in the percent of students who identified vocabulary as being a skill they thought would be improved by playing smartphone games in English from pre- to post-game, the drop was statistically non-significant (p=0.13). The gender spread across each response category was also roughly equivalent to the gender spread for the dataset. Furthermore, there were no significant shifts in the gender spread across each response category from pre- to post-game. These results provide limited evidence that the voice-acted narrative content of LIS significantly altered the perceptions of what linguistic skills could be improved through smartphone gaming, namely that of listening skills.

Figure 6.30 Perception of linguistic improvement from smartphone gaming category breakdown (post-game)

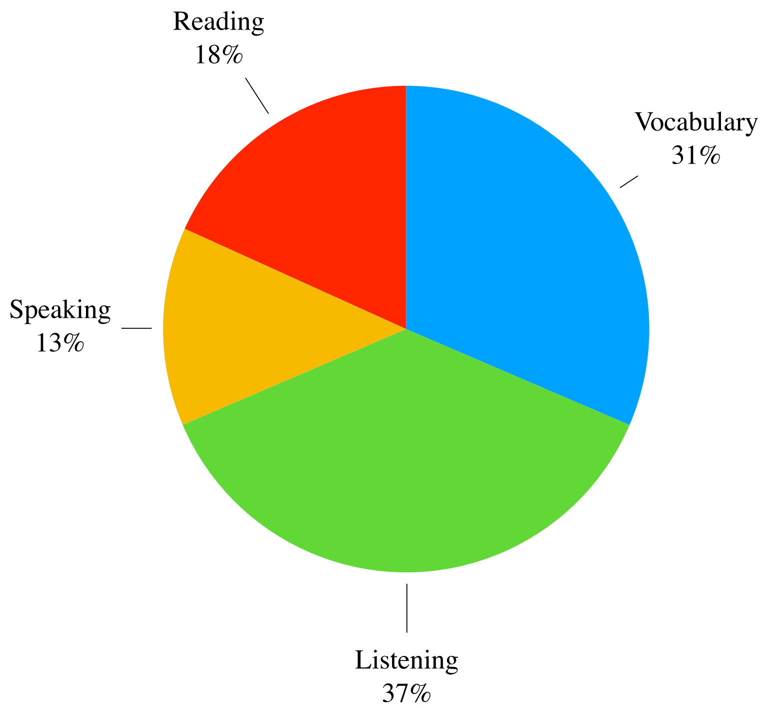
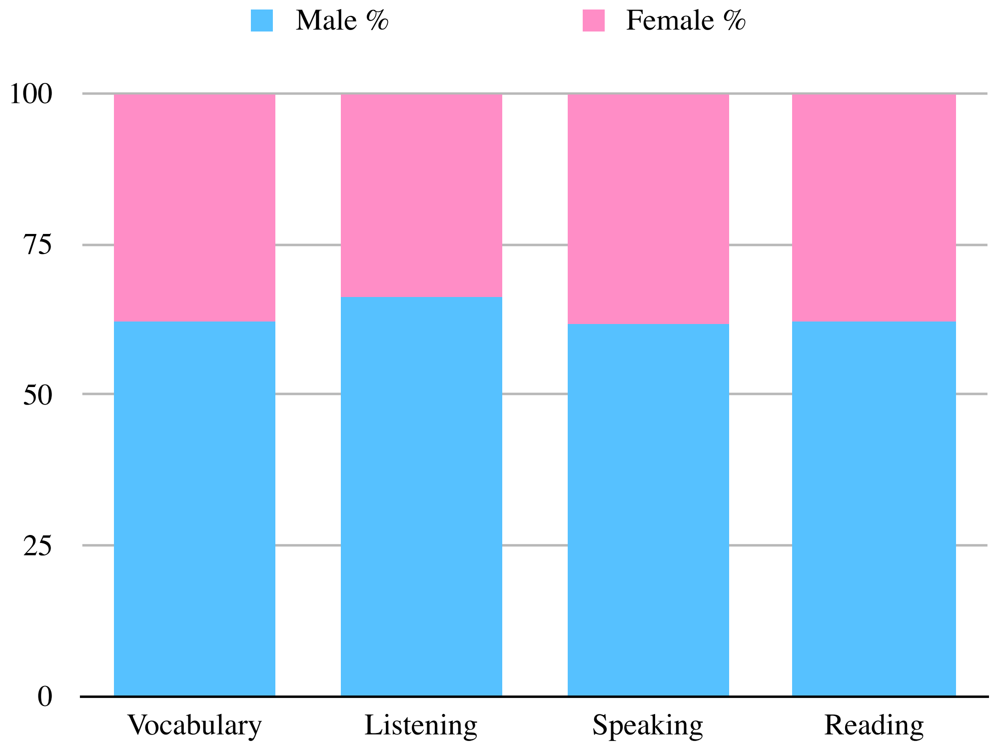


Figure 6.31 Perception of linguistic improvement from smartphone gaming gender breakdown (post-game)



6.7 Analysis of motivation to complete LIS

In order to determine what impact, if any, the desire to continue playing LIS after completing the study program had on vocabulary test scores, a post-game survey question was included to ask students if they were interested in playing through the remaining chapters of the game. After isolating the vocabulary test scores to the two subsequent groups, namely motivated to complete (M) and not motivated to complete (NM), and their respective gender groups, a battery of comparison and correlation analyses were run to identify any significant differences and correlations both between and within each group. The results of the analyses are summarized in Table 6.6. Between-group, two-sample, two-tailed *t*-tests revealed no significant differences between the M and NM groups. Similarly, within-group, two-sample, two-tailed *t*-tests revealed no significant difference between the gender groups in either M or

NM groups. There were, however, a number of significant results from the correlation analyses.

Table 6.6 Analysis of motivation to complete LIS

	Aggregate (*n*=78)	Males (*n*=51)	Females (*n*=27)
Vocabulary test averages and motivation to complete LIS			
Motivation to complete	*n*=50	*n*=32	*n*=18
	54.38	52.99	56.84
No motivation to complete	*n*=28	*n*=19	*n*=9
	57.66	58.54	55.8
Correlation to comprehension tests			
Motivated	*r*-value		
	0.64	0.6	0.098
	p-value		
	0.000098	0.0052	0.77
Not motivated	*r*-value		
	0.43	-0.74	0.99
	p-value		
	0.19	0.056	0.012
Significant gender differences			
Motivated males (MM) to non-motivated males (NMM)	Fisher's *z* (MM)	Fisher's *z* (NMM)	Observed *z*
	0.69	-0.96	2.97
Motivated females (MF) to non-motivated females (NMF)	Fisher's *z* (MF)	Fisher's *z* (NMF)	Observed *z*
	0.10	2.55	2.31
NMM to NMF	Fisher's *z* (NMM)	Fisher's *z* (NMF)	Observed *z*
	-0.96	2.55	3.14
Between-group *t*-tests			
Motivated to not motivated	0.26	0.089	0.86
Within-group, two-sample gender *t*-tests			
Motivated			0.30
Not motivated			0.58

Significant positive correlations between vocabulary and comprehension test scores were observed in both the aggregate (see Figure 6.32) and male (see Figure 6.33) samples in the M group as well as in the female sample of the NM group (see Figure 6.34). Furthermore, all positive correlations observed were found to be strong in the M group aggregate (r=0.64,

p=0.000098), M group males (r=0.6, p=0.0052), and NM group females (r=0.99, p=0.012). In contrast to these positive correlations, a strong negative correlation of near-significance (r=-0.74, p=0.056) was observed in the NM group males (see Figure 6.35). In addition to these significant and near-significant correlations, statistically significant differences between r-values were observed by calculating the observed z-values for compared pairs. The difference in r-values between M group males (r=0.60) and NM group males (r=-0.74) was found to be highly statistically significant (z=2.97, p=0.003). Similarly, a highly statistically significant difference (z=3.14, p=0.0017) was observed between NM males (r=-0.74) and NM group females (r=0.99). Finally, although the correlation coefficient in the M group female sample was found to be non-significant and negligible (r=0.098, p=0.77) (see Figure 6.36), the difference between the NM group female values was found to be significant (z=2.31, p=0.021). A limited interpretation of these results will be discussed in the following section.

Figure 6.32 Motivation to complete LIS correlation analysis 1

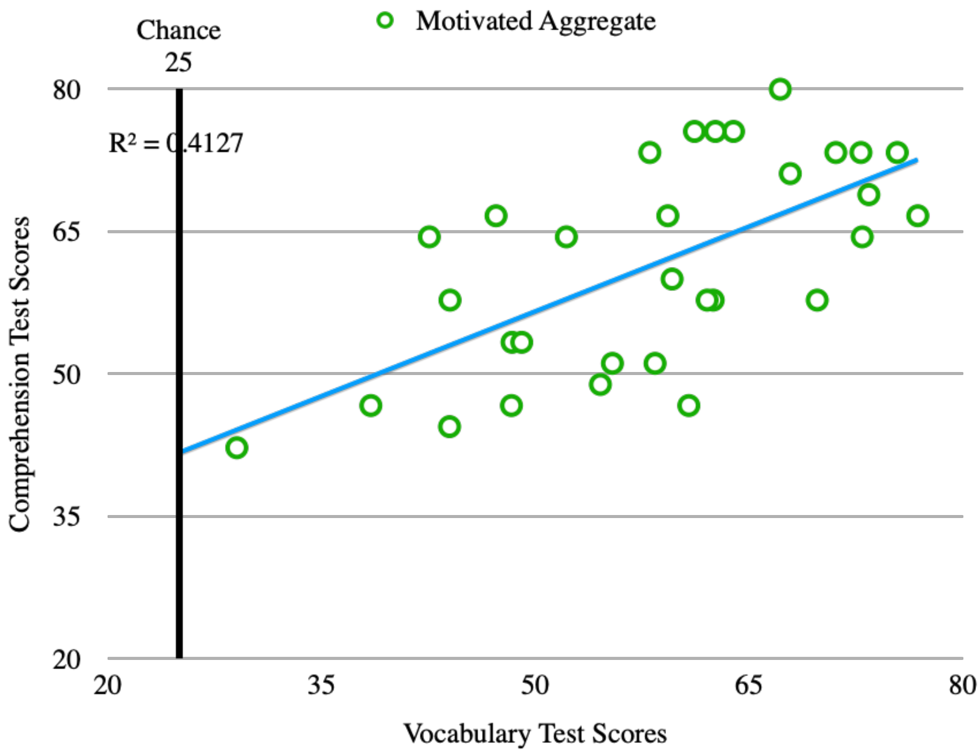


Figure 6.33 Motivation to complete LIS correlation analysis 2

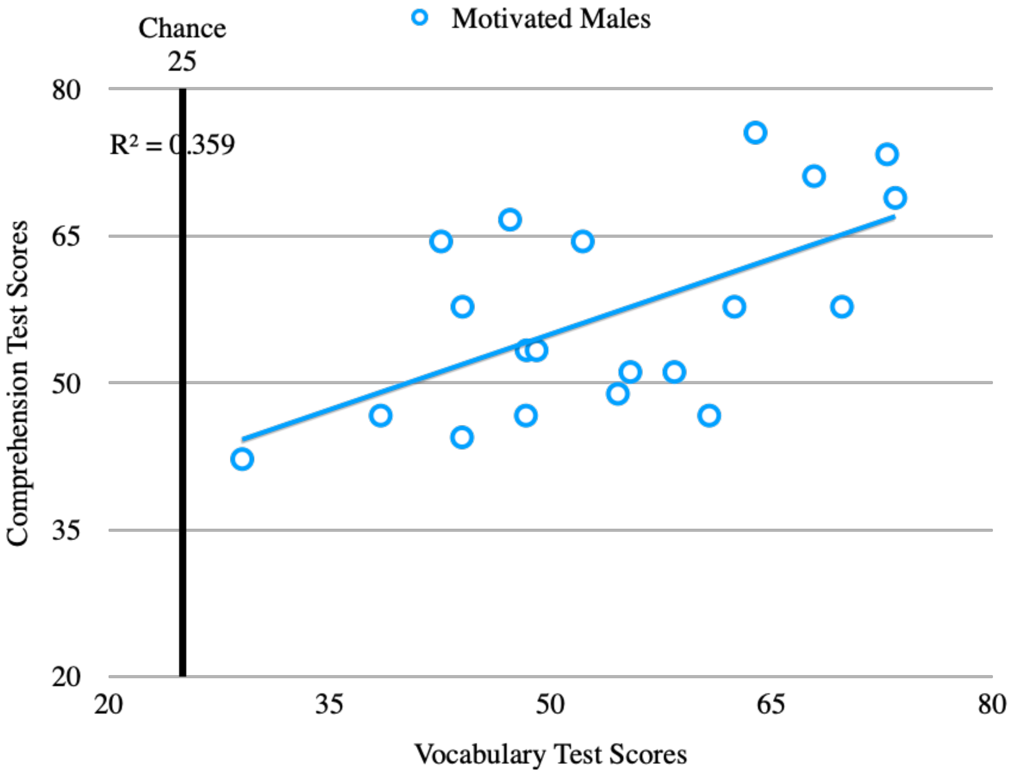


Figure 6.34 Motivation to complete LIS correlation analysis 3

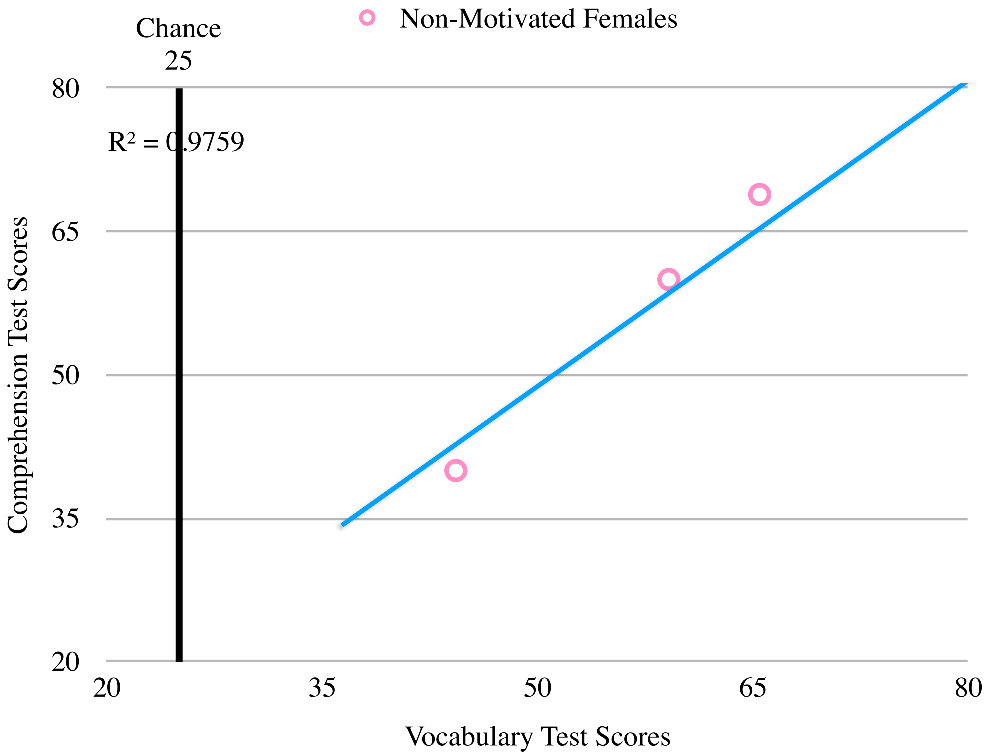


Figure 6.35 Motivation to complete LIS correlation analysis 4

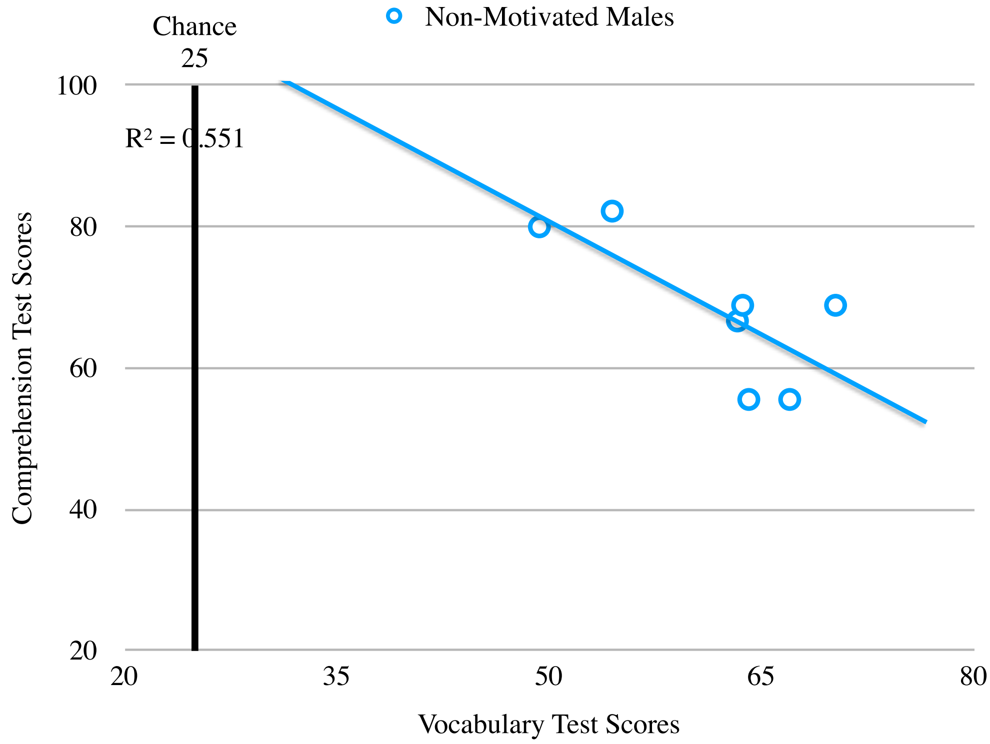
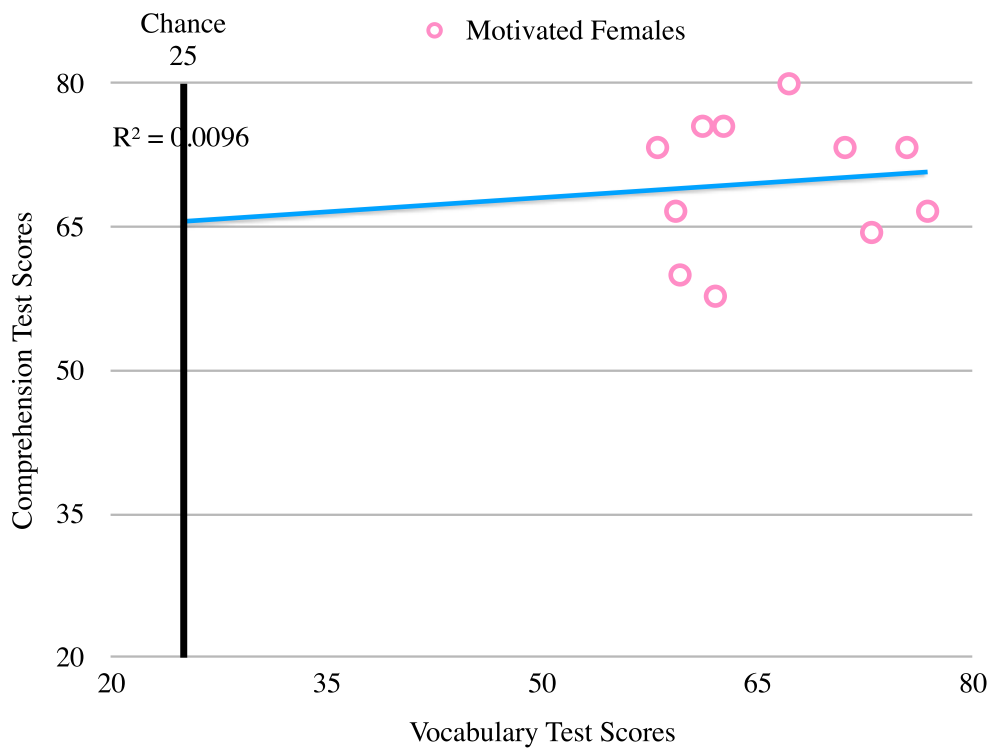


Figure 6.36 Motivation to complete LIS correlation analysis 5



On the whole, both aggregate samples in the M and NM groups demonstrated approximately the same correlation between vocabulary test scores and comprehension test scores as was observed in the full modified LG group aggregate sample. With that being said, the two most anomalous results were the statistically significant negative correlation in the NM group male sample (see Figure 6.35) and the non-significant, negligible correlation in the M group female sample (see Figure 6.36). The vocabulary and comprehension scores in the M group female sample are largely clustered in the upper ranges of test averages for *x* and *y* axes, thus suggesting that the females in the dataset who self-identified an interest in the game may have had higher levels of English proficiency relative to other groups in the dataset. The negative correlation for the non-motivated males is, however, more difficult to explain and largely lies outside the scope of the present analysis to confidently interpret.

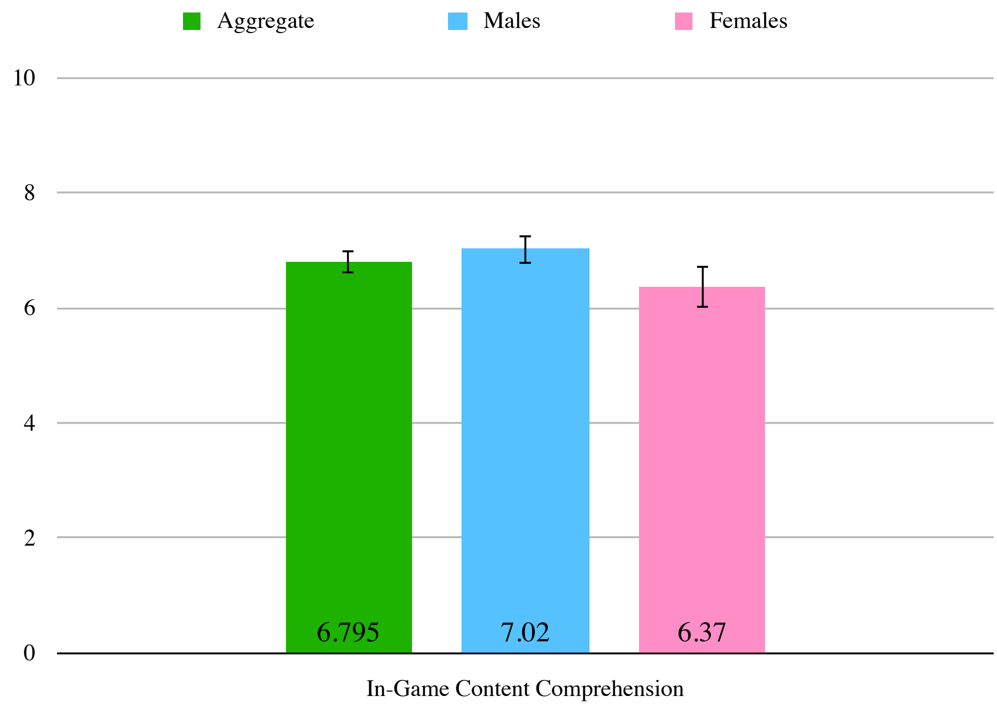
One argument that may be put forward is that non-motivated males with high comprehension test scores and low vocabulary test scores had higher levels of English proficiency which allowed them to complete the in-game tasks quicker than average, which would, in turn, reduce their exposure time to the retention-building game mechanics for vocabulary acquisition. Conversely, the non-motivated males with lower levels of English proficiency may have had more difficulty comprehending in-game narrative content, which, in turn, would arguably have resulted in them taking more time to complete in-game tasks and exposed them to more retention-building game mechanics for vocabulary acquisition. This interpretation would imply, however, that all LG group participants with lower levels of English proficiency and who needed more time to complete the in-game tasks should display higher vocabulary test scores, a supposition which is not reflected in any aggregate data samples analyzed in the present study.

Another possible interpretation is that data from the relatively low sample size in the non-motivated male group (n=19) are being skewed by non-motivated, multiple play-through students (LG85 and LG96) and a student with prior LIS gaming experience in English (LG69). However, when removing these three students' data from the dataset, the negative correlation actually becomes more pronounced (r=-0.88, p=0.12). Ultimately more data would be required to address the source of the inverse relationship between in-game narrative content comprehension and vocabulary test scores in non-motivated males. In general, the data seem to suggest that motivation to complete LIS (or lack thereof) affects the relationship between in-game comprehension and subsequent vocabulary scores.

6.8 Analysis of in-game content comprehension

The following analysis of the students' perceptions of in-game content comprehension was based on data gathered from the post-game survey. Students were asked to rate how well they understood the content of LIS on a 10-point Likert scale (1=not at all, 10=very well), the results of which are summarized in Figure 6.37. Within-group, two-sample, two-tailed t-tests were used to test for significant differences in Likert values between the gender groups, for which no significant differences were observed.

Figure 6.37 Perception of in-game content comprehension Likert responses



Correlation analyses were also conducted for the perception of in-game content comprehension survey question to analyze potential correlations with other Likert scale survey items as well as with vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.7. Statistically significant correlations were revealed between the perception of in-game content comprehension values and the perception of English growth values. Although the correlation between the perception of in-game content comprehension values and comprehension test score values were negligible and non-significant, a moderately significant difference between the *r*-values in the gender groups was observed. The following section will detail both significant and moderately significant findings.

Table 6.7 Correlation analysis of in-game content comprehension

	Aggregate (n=78)		Males (n=51)		Females (n=27)	
Categories	r-value	p-value	r-value	p-value	r-value	p-value
Vocabulary scores	0.079	0.49	0.17	0.24	-0.017	0.93
Comprehension scores	-0.0044	0.98	0.23	0.26	-0.22	0.42
Post-game smartphone engagement	0.076	0.51	0.047	0.74	0.11	0.58
Post-game perception of smartphone gaming efficacy	0.20	0.078	0.11	0.45	0.34	0.08
Content difficulty	-0.10	0.37	-0.13	0.35	-0.043	0.83
Perception of English growth	0.23	0.047	0.36	0.0094	-0.0053	0.98
Categories with significant gender differences	Observed z-value		Fisher's z (males)		Fisher's z (females)	
Comprehension scores	1.83		0.23		-0.23	

A moderately weak correlation (r=0.23, p=0.047) was observed between the perception of in-game content comprehension values and the perception of English growth values in the aggregate sample (see Figure 6.38) and a moderate correlation (r=0.36, p=0.0094) was observed in the male sample for the same values (see Figure 6.39). Both correlations are positive, indicating that on the whole, students who better understood the in-game content were more likely to perceive that their English benefited as a result of playing the game. However, this correlation is not only far more pronounced in the male sample, but the correlation is altogether negligible and non-significant in the female sample (*r*=-0.0053, *p*=0.98). These results lend themselves to the argument that a potentially significant gender difference regarding the comprehension of in-game content may affect how some male students in the present study judge their English development, and in turn the value they place on the curriculum itself.

Figure 6.38 Perception of in-game content comprehension correlation analysis 1

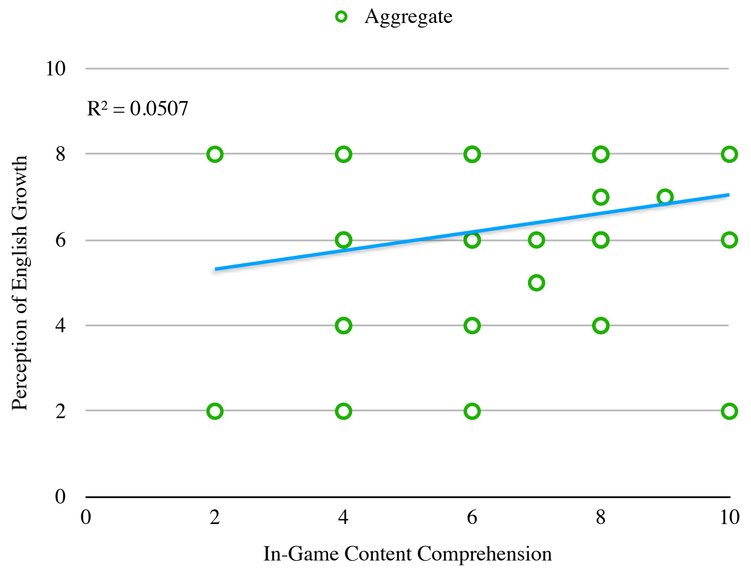
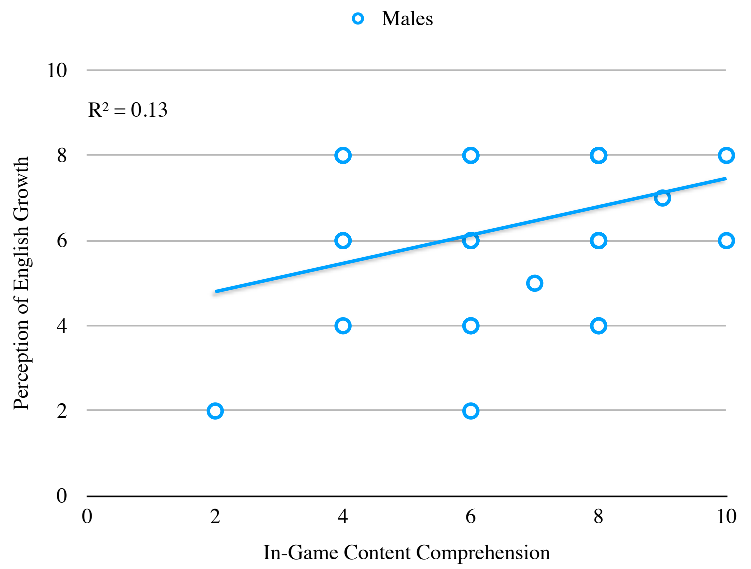


Figure 6.39 Perception of in-game content comprehension correlation analysis 2



The correlation between the perception of in-game content comprehension and comprehension test scores, while negligible and non-significant in the aggregate sample (*r*=-

0.0044, p=0.98), this near-zero correlation result in the aggregate sample is largely due to a cancelling out effect from a positive correlation (r=0.23, p=0.26) in the male sample (see Figure 6.40) and a negative correlation (r=-0.22, p=0.42) in the female sample (see Figure 6.41). When comparing these inverse correlations between the gender groups with Fisher's z transformations and then performing an observed z calculation, a near significant result (z=1.83, p=0.067) was observed. The moderately significant difference between the two gender groups may support a line of argument postulating that there is some fundamental difference between how males and females define "comprehension of in-game content." More specifically, considering the significantly higher comprehension test averages from the female sample, the result of a negative correlation between self-reported in-game comprehension values and comprehension test scores is highly notable. One interpretation for the inverse comprehension correlation in the female sample is that a small portion of the female sample is qualitatively basing the concept of "in-game comprehension" on more technical aspects (e.g., game controls) of the game rather than the comprehension of narrative content. The following section will provide limited evidence to support this data interpretation.

Figure 6.40 Perception of in-game content comprehension correlation analysis 3

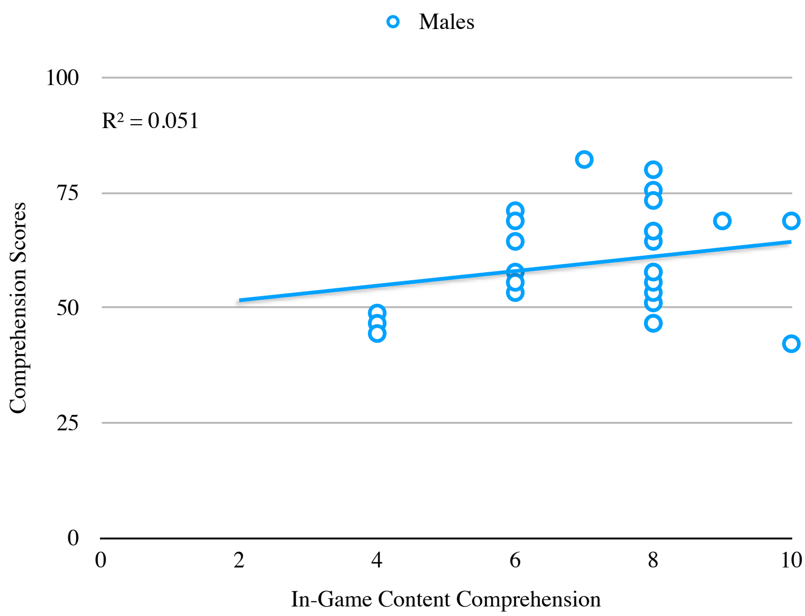
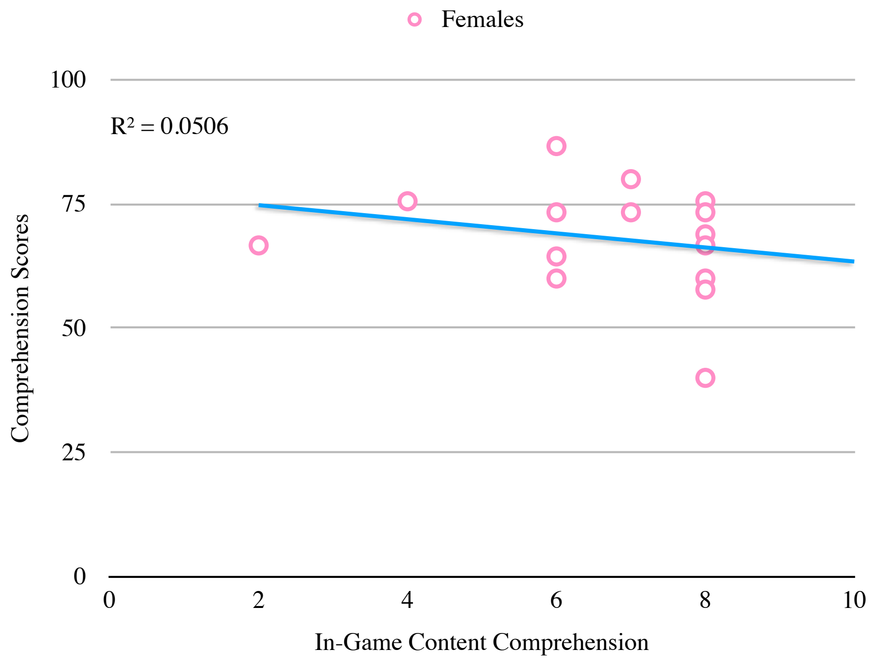


Figure 6.41 Perception of in-game content comprehension correlation analysis 4



In order to analyze which specific aspects of the in-game content were difficult to understand, a post-game survey question asked students to list anything about the game's content that they found to be particularly difficult to comprehend. The results of the post-game survey question were subjected to recursive analysis to identify common response categories, which were then subjected to quantitative analysis. There were a total of six response categories identified and these are summarized in Figure 6.42. The gender breakdown for each response category is summarized in Figure 6.43. Compared to the gender percentage for the total dataset (males=65.39%, females=34.62%), there were a number of categories with significant gender-group skews. Response categories with higher male to female ratios included: 1) entertainment value (e.g., the game's gimmick) (males=100%), 2) character relationships (e.g., the roles of each character) (males=77.78%), and 4) story content/game mechanics (e.g., how to progress in the game) (males=77.78%). Response categories with higher female to male ratios included: 3) dialogue (e.g., the speed of the spoken language) (females=51.52%) and 5) game controls (e.g., difficulty in performing certain in-game actions such as time reversal) (females=60%). In line with the argument presented above, a greater proportion of females qualitatively reported difficulty in comprehending certain aspects of the game such as the in-game controls, which arguably have little impact in the comprehension of the in-game narrative content.

Figure 6.42 In-game content comprehension response category breakdown

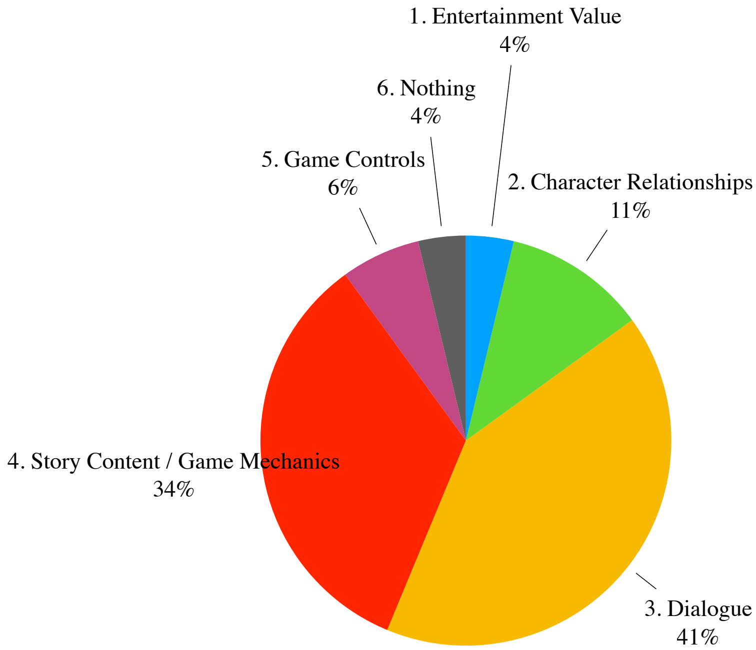
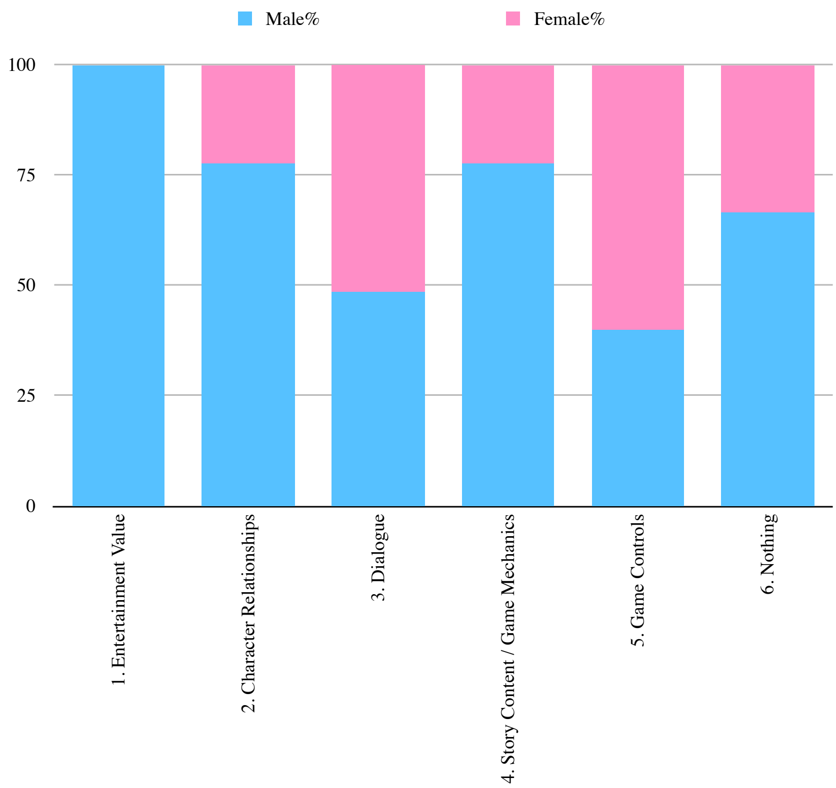


Figure 6.43 In-game content comprehension gender breakdown



6.9 Analysis of post-game smartphone engagement

The following analysis of the students' intention for engaging with smartphone gaming after having participated in the present study program was based on data gathered in the post-game survey. Students were asked to rate how likely they will be to continue playing smartphone games after the study program on a 10-point Likert scale (1=not at all likely, 10=very likely), the results of which were summarized in Figure 6.2 in the pre-game smartphone engagement analysis. Within-group, two-sample, two-tailed *t*-tests were used to test for significant differences in Likert values between the gender groups, for which no significant differences were observed. Between-group *t*-tests between pre- and post-game values revealed no significant differences. Correlation analyses were also conducted for the post-game smartphone engagement survey question to analyze potential correlations with other Likert scale survey questions as well as with vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.8. Statistically significant correlations were revealed between the post-game smartphone engagement values and the perception of English growth values. No other significant correlations or differences in gender group *r*-values were observed. The following section will detail the statistically significant findings discussed above.

Table 6.8 Correlation analysis of post-game smartphone engagement

	Aggregate (n=78)		Males (n=51)		Females (n=27)	
Categories	r-value	p-value	r-value	p-value	r-value	p-value
Vocabulary scores	0.13	0.24	0.11	0.45	0.18	0.38
Comprehension scores	0.07	0.67	-0.04	0.86	0.23	0.41
Post-game perception of smartphone gaming efficacy	0.12	0.31	0.20	0.15	-0.065	0.75
Content difficulty	-0.062	0.59	0.08	0.56	-0.29	0.15
Perception of English growth	0.31	0.0057	0.24	0.091	0.41	0.033

Moderate correlations between post-game smartphone engagement values and perception of English growth values were observed in the aggregate sample (r=0.31,

p=0.0057) (see Figure 6.44) and female sample (r=0.41, p=0.033) (see Figure 6.45). Although a similar positive correlation was observed in the male sample for the same category, the correlation (r=0.24, p=0.091) was less pronounced than in the female sample and above the alpha threshold for statistical significance. When comparing the significant correlations to the perception of English growth values to the values from both pre- and post-game smartphone engagement datasets (see Table 6.1), similar positive correlations were observed for both the aggregate and male samples. However, in the female sample, the correlation between smartphone gaming engagement values and the perception of English growth values switched from a pre-game negative correlation (r=-0.28, p=0.15) (see Table 6.1) to a statistically significant positive correlation (see Figure 6.45), meaning that a certain amount of females had to, in course of the study program, change their opinions about playing smartphone games and that this shift is related to how much that sample perceived their English skills to be improved by playing LIS.

Figure 6.44 Post-game smartphone engagement correlation analysis 1

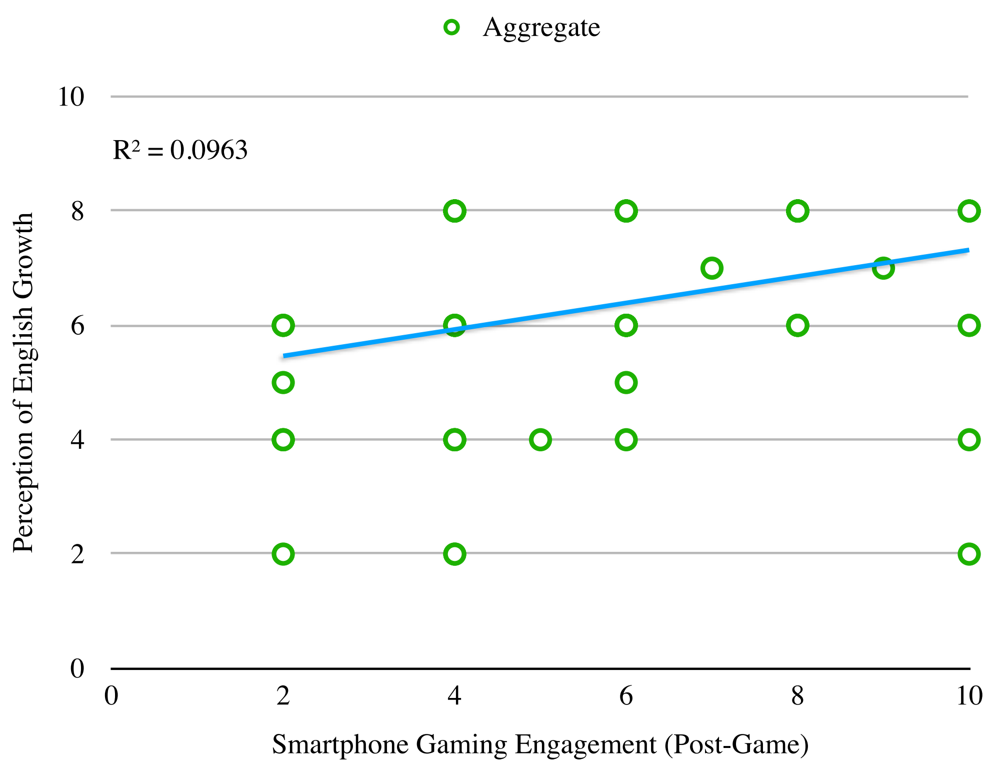
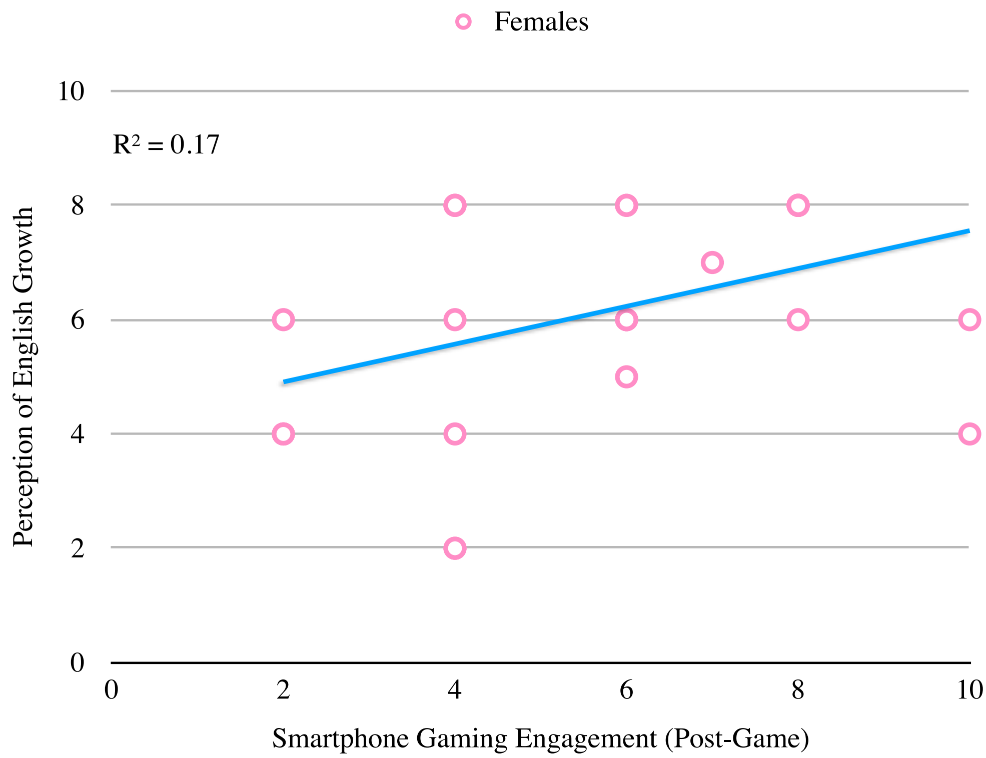


Figure 6.45 Post-game smartphone engagement correlation analysis 2



Upon analyzing the source of this correlation inversion from pre- to post-game, a significant shift in the smartphone gaming engagement in the entire female sample was ruled out. The pre-game smartphone engagement correlation analysis revealed a moderate, statistically significant correlation between pre-game and post-game smartphone engagement values in the female sample (r=0.38, p=0.05) (see Figure 6.5). Furthermore, the averages of the Likert scale responses revealed no significant shifts in the engagement with smartphone games from pre-game (see Figure 6.1) to post-game (see Figure 6.2). Therefore, the source of the shift was determined to be coming from a switch in some of the extreme outliers of the female sample, which would explain the correlation shift to perception of English growth from a pre-game negative value to a post-game positive value while still preserving the positive correlations between pre- and post-game smartphone engagement and equivalent pre- and post-game Likert response averages.

With this data interpretation in mind, the females who engaged with smartphone games the least prior to participating in this study were the most likely to have perceived that playing LIS improved their English and were, in turn, more likely to consider playing smartphone games in the future. Likewise, females who engaged with smartphone games the most prior to participating in this study were less likely to have perceived that playing LIS improved their English and were, in turn, less likely to consider playing smartphone games in the future. In addition to the shift in correlation between smartphone gaming engagement and perception of English growth from pre- to post-game in the female sample, other pre- to post-game shifts in correlational relationships were observed with the vocabulary scores, comprehension scores, in-game content comprehension, perception of in-game English difficulty, and pre-/post-game perception of smartphone gaming efficacy. The correlational shifts in each Likert response category will be discussed in detail below.

Negligible, non-significant negative correlations were observed between the pre-game smartphone gaming engagement values and the vocabulary test scores in the aggregate (r=-0.16, p=0.17), male (r=-0.2, p=0.16), and female samples (r=-0.09, p=0.66). There were shifts to negligible, non-significant positive correlations in the post-game smartphone gaming engagement values in the aggregate (r=0.13, p=0.24), male (r=0.11, p=0.45), and female samples (r=0.18, p=0.38). When comparing the pre-/post-game smartphone engagement values to the vocabulary test scores in the aggregate (see Figure 6.46 and Figure 6.47), male (see Figure 6.48 and Figure 6.49), and female samples (see Figure 6.50 and Figure 6.51), the negative to positive shifts become apparent. Although the above-stated correlations are negligible and non-significant, the degree of the inversion between the correlational relationships is notable and provides limited evidence that the LIS gaming intervention in conjunction with performance on the vocabulary tests may have affected how willing some students were to engage with smartphone games after having participated in the study. More specifically, students (both male and female) who performed well on the vocabulary tests were more likely to report that they had the intention to play more smartphone games on conclusion of the study program.

Figure 6.46 Post-game smartphone engagement correlation analysis 3

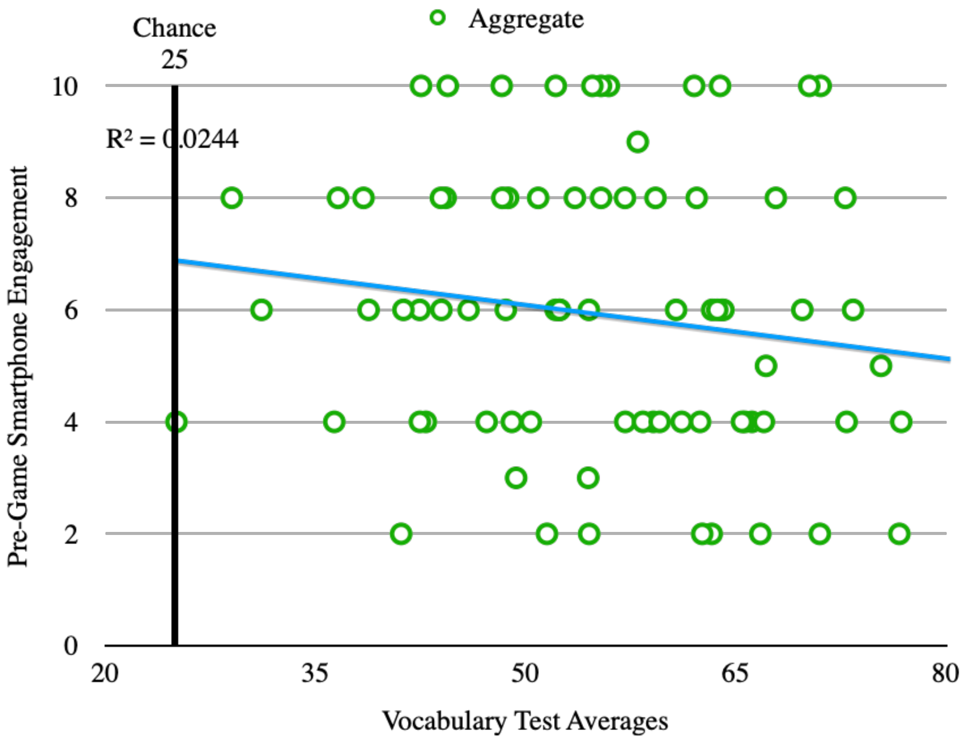


Figure 6.47 Post-game smartphone engagement correlation analysis 4

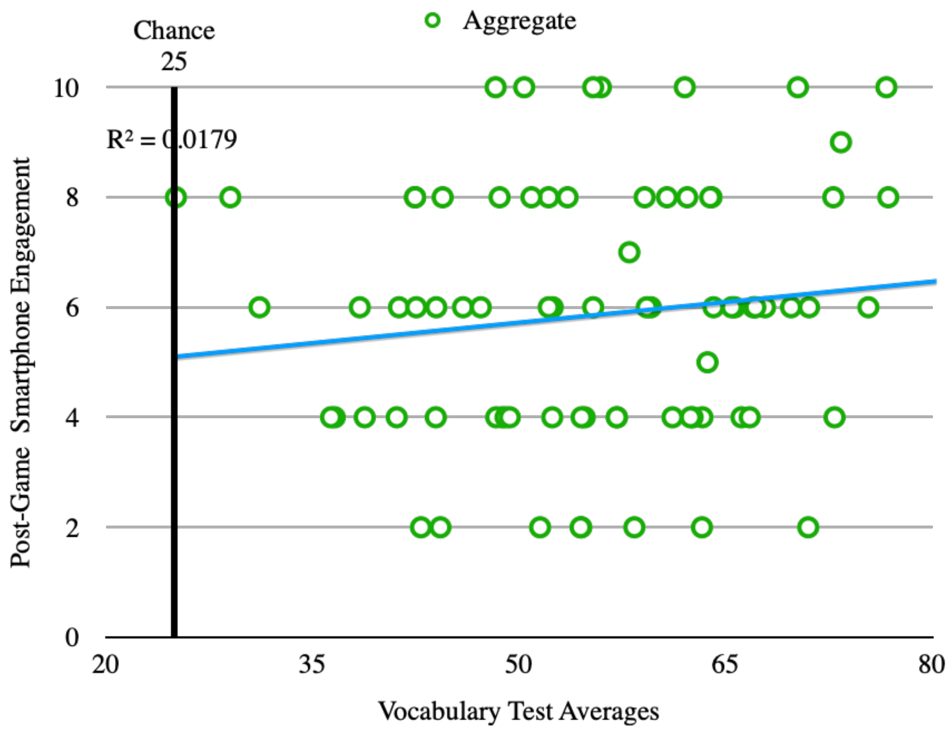


Figure 6.48 Post-game smartphone engagement correlation analysis 5

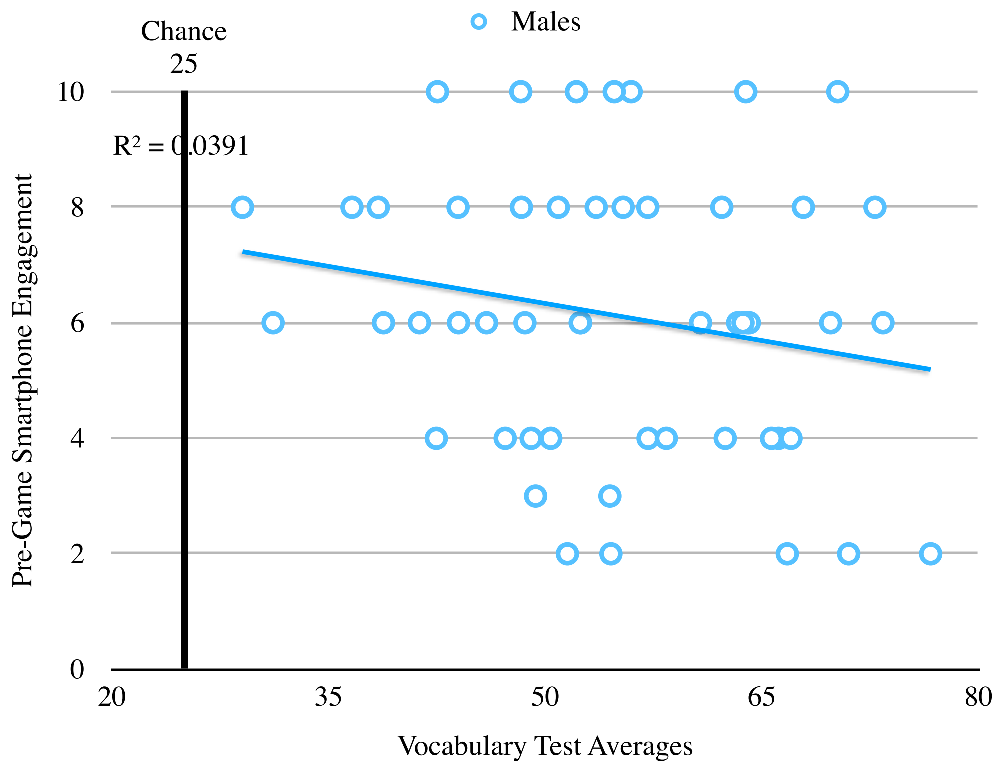


Figure 6.49 Post-game smartphone engagement correlation analysis 6

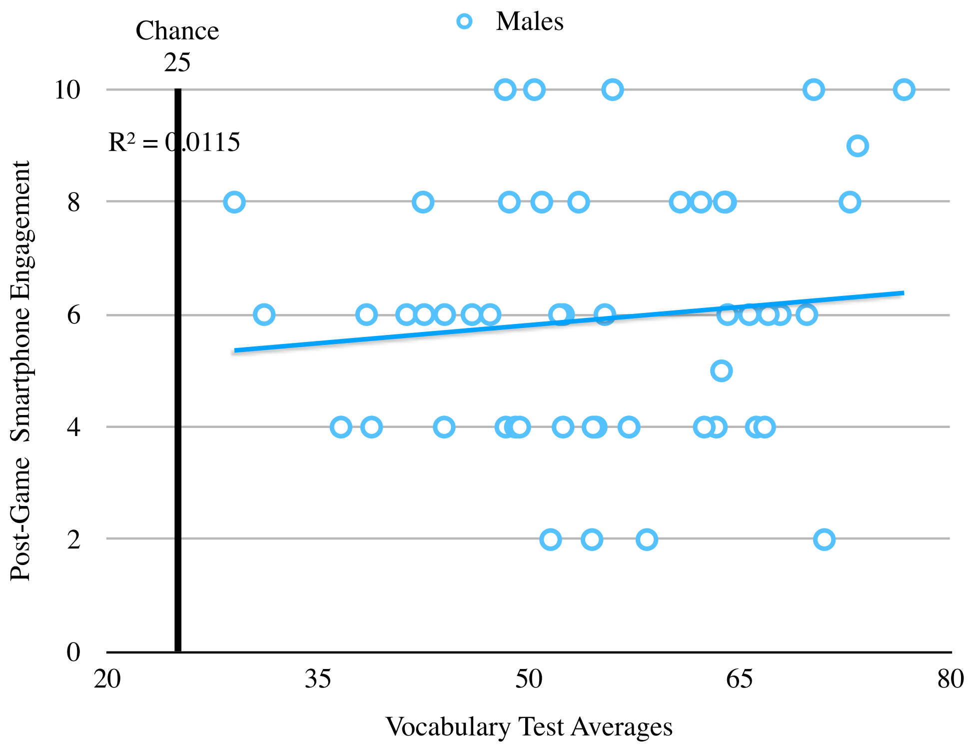


Figure 6.50 Post-game smartphone engagement correlation analysis 7

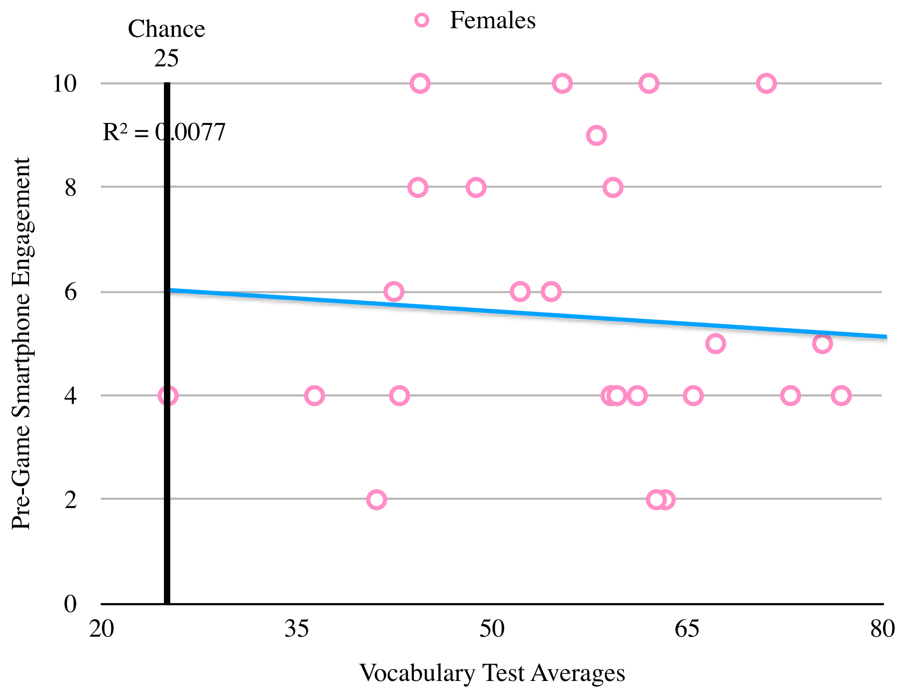
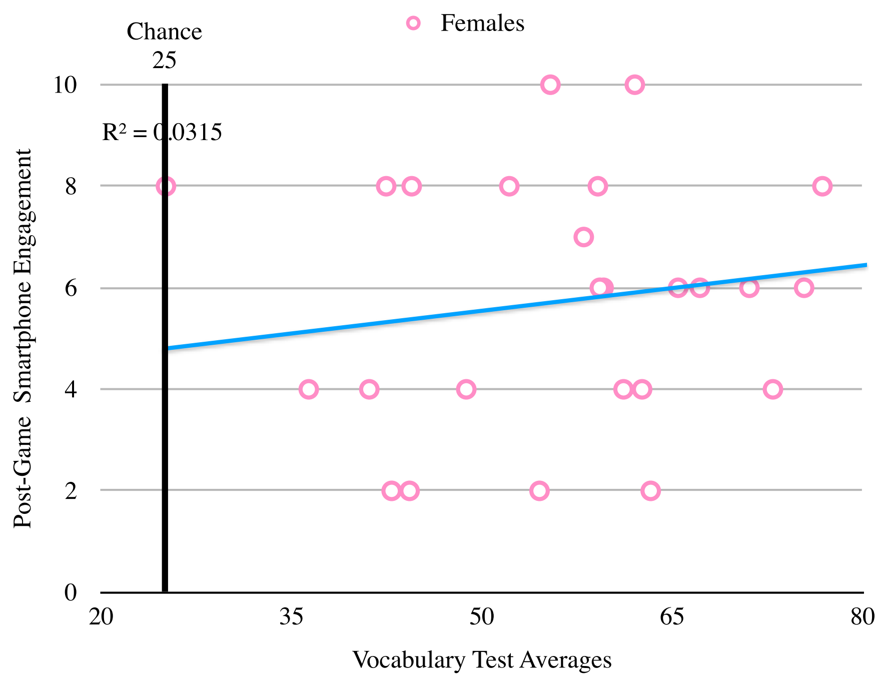


Figure 6.51 Post-game smartphone engagement correlation analysis 8



Negligible, non-significant negative correlations were also observed between the pre-game smartphone gaming engagement values and the comprehension test scores in the aggregate (r=-0.23 p=0.15) and female samples (r=-0.37, p=0.17). In the post-game smartphone gaming engagement values, there was a shift to a negligible, non-significant positive correlation in the aggregate (r=0.07, p=0.67) and female samples (r=0.23, p=0.41). No significant shift was observed in the male sample from pre- to post-game smartphone engagement values in the correlation to comprehension test scores. A comparison of the negative to positive correlational shifts in the aggregate sample (see Figure 6.52 and Figure 6.53) and female sample (see Figure 6.54 and Figure 6.55) are summarized below. Just as with the vocabulary score correlations described above, the performance on the comprehension test was correlated to shifts in the intention to engage with smartphone games upon conclusion of the study program. In contrast to the vocabulary test correlations, however, the source of the correlational shift was localized to the female population. These data may suggest that although the post-game intentions to engage with smartphone games for a certain number of students in both gender groups were affected by the LIS gaming intervention, the female group may place more emphasis on the comprehension of in-game content than the males who participated in this study.

Figure 6.52 Post-game smartphone engagement correlation analysis 9

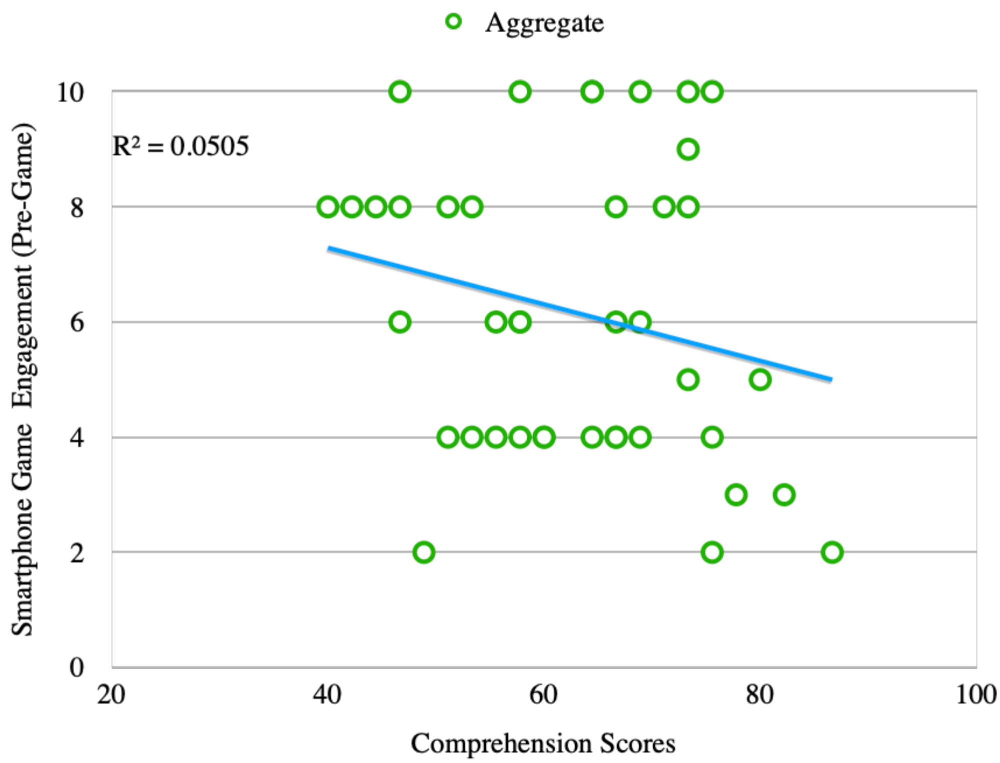


Figure 6.53 Post-game smartphone engagement correlation analysis 10

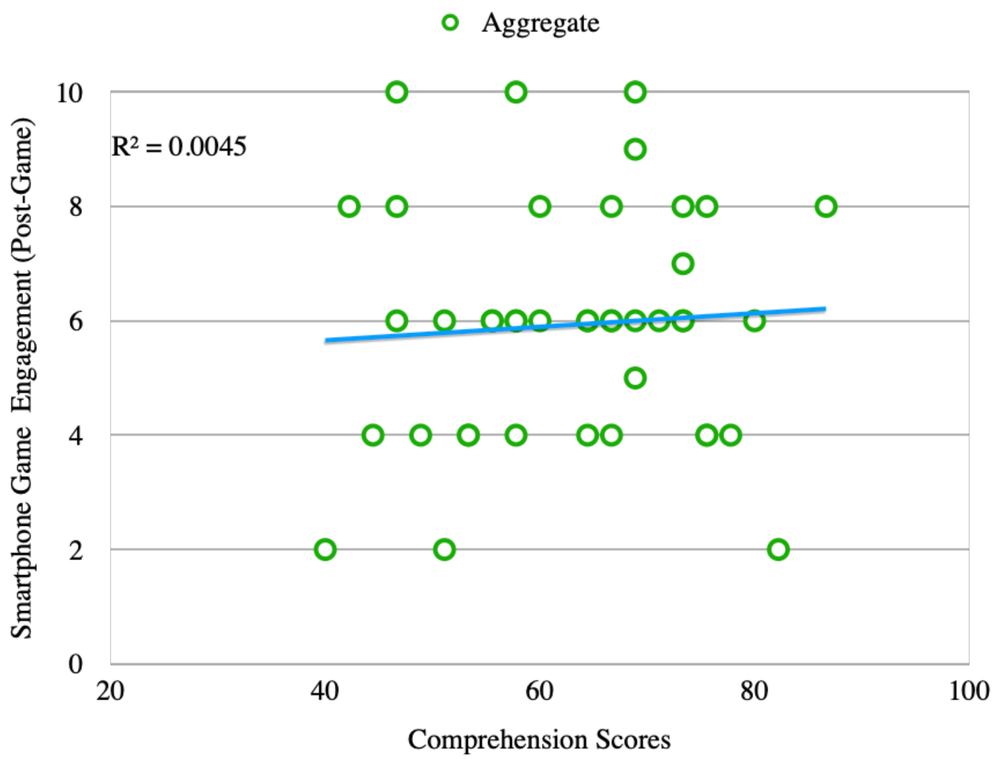


Figure 6.54 Post-game smartphone engagement correlation analysis 11

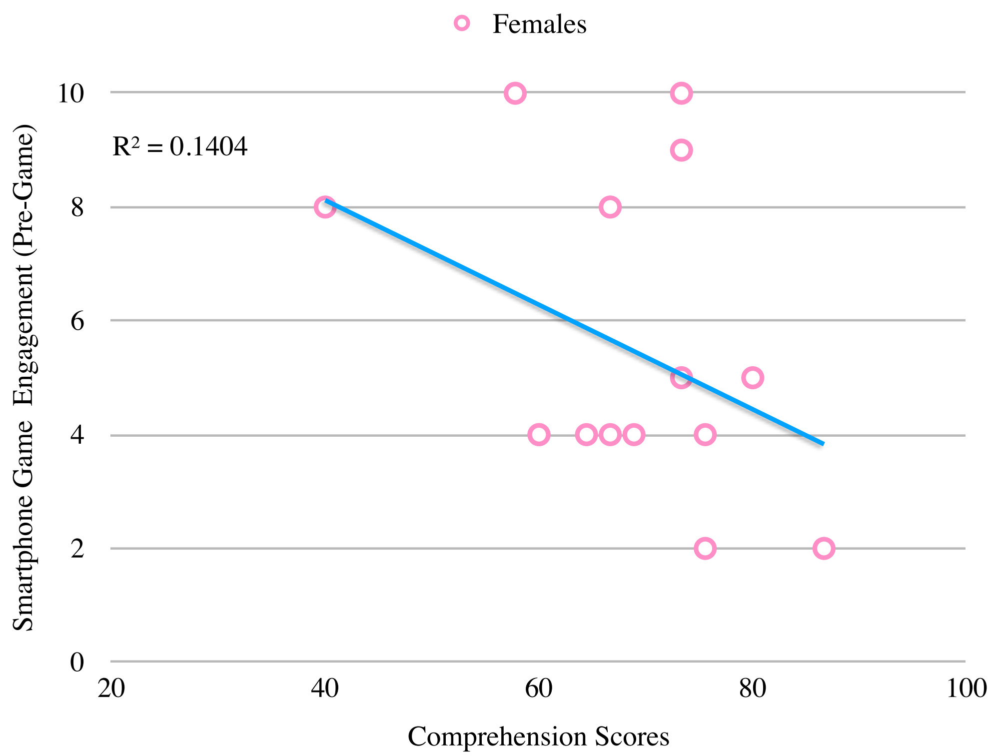
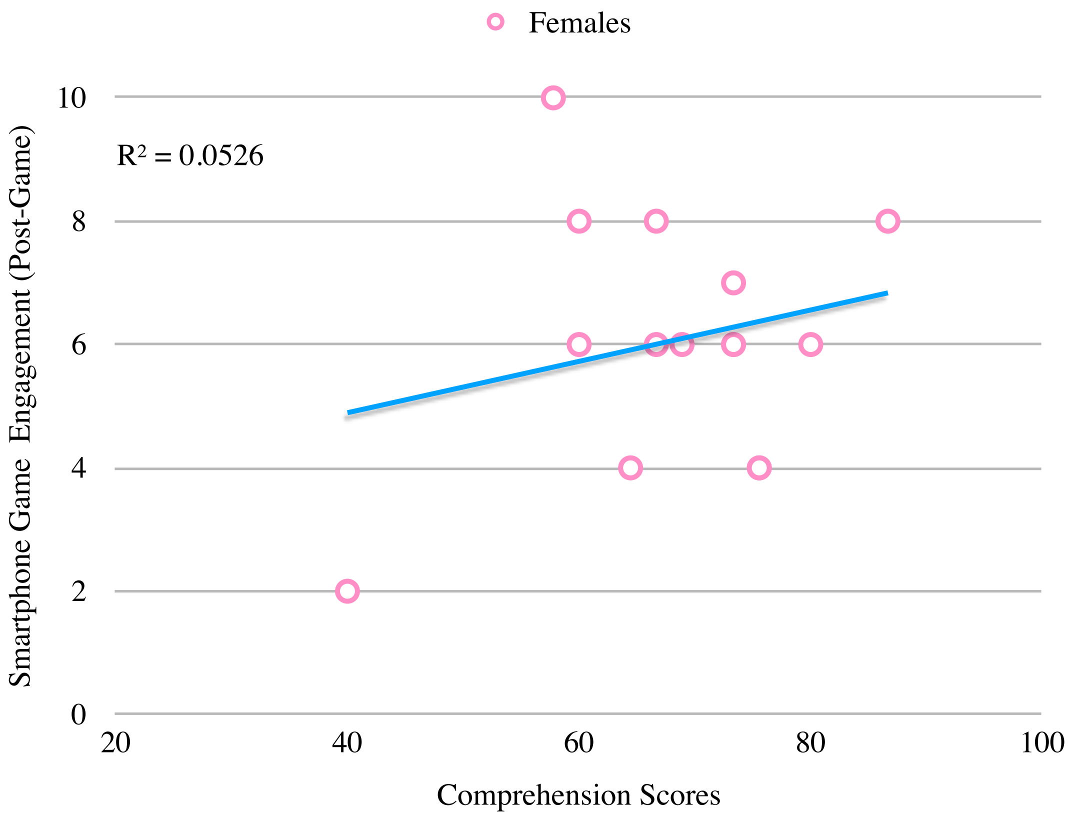


Figure 6.55 Post-game smartphone engagement correlation analysis 12



6.10 Analysis of post-game perception of smartphone gaming efficacy

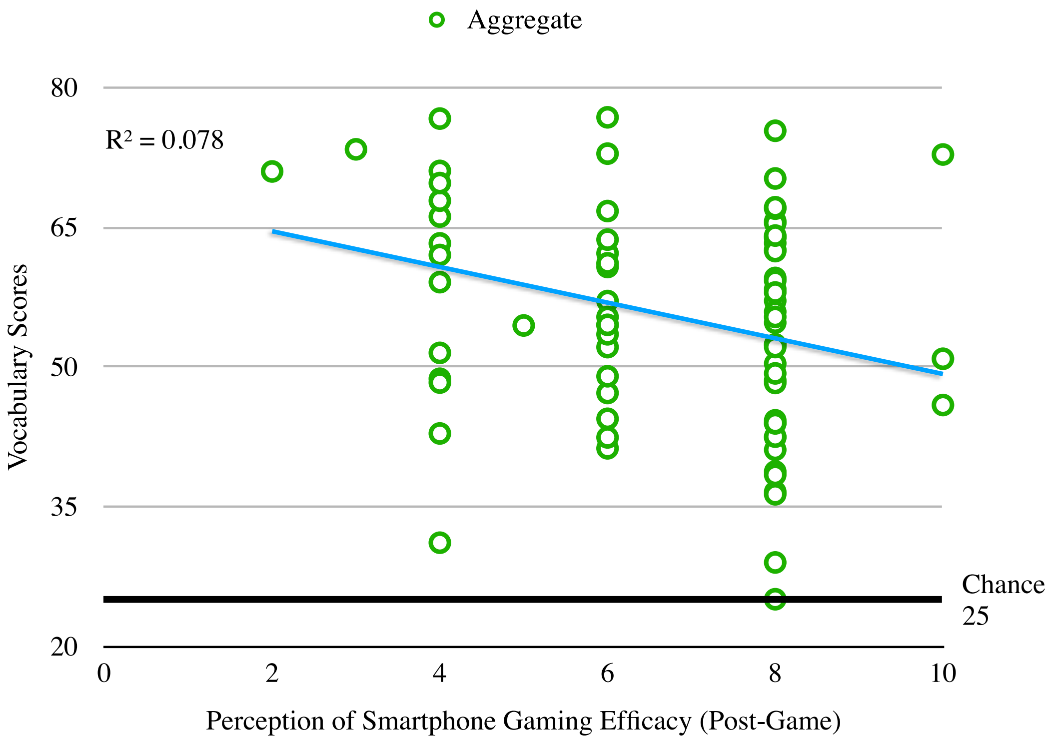
The following analysis of the students' perception of the efficacy of using smartphone gaming as a language learning tool after having participated in the present study program was based on data gathered in the post-game survey. Students were asked to rate how efficacious they perceived smartphone games to be as tools to learn English on a 10-point Likert scale (1=not at all, 10=very efficacious), the results of which were summarized in Figure 6.18 in the pre-game perception of smartphone efficacy analysis. Within-group, two-sample, two-tailed *t*-tests were used to test for significant differences in Likert values between the gender groups, for which no significant differences were observed. Between-group *t*-tests between pre- and post-game values also revealed no significant differences. Correlation analyses were also conducted for the post-game smartphone engagement survey item to analyze potential correlations with other Likert scale survey items as well as with vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.9. Statistically significant correlations to the post-game perception of smartphone gaming efficacy values were revealed between the vocabulary test scores, the content difficulty of the in-game English, and the perception of English growth values. The following section will detail the statistically significant findings discussed above.

Table 6.9 Correlation analysis of the post-game perception of smartphone gaming efficacy

	Aggregate (*n*=78)		Males (*n*=51)		Females (*n*=27)	
Categories	*r*-value	*p*-value	*r*-value	*p*-value	*r*-value	*p*-value
Vocabulary scores	-0.28	0.013	-0.26	0.07	-0.32	0.11
Comprehension scores	-0.15	0.35	-0.13	0.52	-0.092	0.75
Content difficulty	0.28	0.013	0.15	0.29	0.54	0.0038
Perception of English growth	0.34	0.002	0.47	0.00051	0.12	0.55
Categories with significant gender differences	Observed *z*-value		Fisher's *z* (males)		Fisher's *z* (females)	
Content difficulty	1.80		0.15		0.60	

A moderately weak negative correlation (r=-0.28, p=0.013) to vocabulary scores was observed in the aggregate sample (see Figure 6.56). This significant negative correlation between the post-game perception of smartphone gaming efficacy as a language learning tool and vocabulary test scores is approximately the same compared with the pre-game perception of smartphone gaming efficacy values (see Figure 6.19). These negative correlations suggest that for students who performed worse on the vocabulary tests, they were slightly more likely to perceive smartphone games as efficacious language learning tools. Furthermore, as the pre-game and post-game smartphone gaming efficacy correlations to vocabulary test scores are nearly identical, the gaming intervention of LIS presented in this study arguably did not have any impact on this correlational relationship.

Figure 6.56 Post-game smartphone efficacy correlation analysis 1



Although there were no significant changes in the correlations between the pre-/post-game smartphone gaming efficacy values and vocabulary test scores, there were significant shifts in correlation between the post-game smartphone gaming efficacy values and comprehension scores relative to the pre-game correlation values (see Figure 6.21, Figure 6.22, and Figure 6.23). The correlation between the pre-game perception of smartphone gaming efficacy values and comprehension test scores revealed moderate to strong, statistically significant negative relationships (i.e., the more efficacious smartphone games were perceived to be as language learning tools, the worse the comprehension test scores) in the aggregate (r=-044, p=0.0035), male (r=-0.4, p=0.040), and female samples (r=-0.67, p=0.0063). These pre-game negative correlations, however, were largely nullified in the post-game values, especially in the female sample. The correlation between the post-test smartphone gaming efficacy values and comprehension test scores were negligible, non-significant negative correlations in the aggregate (r=-0.15, p=0.35), male (r=-0.13, p=0.52), and female samples (r=-0.092, p=0.75). The large shifts in correlation values from pre- to post-game may suggest that the LIS gaming intervention used in the present study affected this correlational relationship between the perception of smartphone gaming as an efficacious language learning tool and overall performance on the comprehension tests. In particular, a certain number of participants (especially from the female sample), who performed well on the comprehension tests as a result of comprehending the content of LIS arguably underwent a change of opinion in how efficacious smartphone games could be as language learning tools.

Another correlational relationship that shifted significantly from pre- to post-game is the correlation between the perception of smartphone gaming efficacy values and the content comprehension values. Whereas the correlations between the pre-game perception of smartphone gaming efficacy values and the in-game content comprehension values were negligible in the aggregate (r=-0.0017, p=0.99), male (r=0.037, p=0.8), and female samples (r=-0.13, p=0.52), the correlations to the post-game perception of smartphone gaming efficacy values approached statistical significance in the aggregate (r=0.20, p=0.078) and female samples (r=0.34, p=0.08), but remained negligible and non-significant in the male sample (r=0.11, p=0.45). Comparing the pre-/post-game aggregate (see Figure 6.57 and Figure 6.58), male (see Figure 6.59 and Figure 6.60), and female sample correlations (see Figure 6.61 and Figure 6.62) side by side, the significant shift in females' perception of the efficacy of smartphone gaming as it relates to how much of the in-game content was comprehended becomes apparent. The data seem to suggest that a significant number of females changed their opinion regarding the efficacy of smartphone gaming as a language learning tool. Specifically, females who comprehended the in-game content were more likely to perceive smartphone games as efficacious language learning tools after having participated in the present study program. These correlation analysis results add weight to the argument outlined in the previous section (i.e., the correlation to comprehension test scores) that females in this sample regard the comprehension of in-game content as a vital component to perceive smartphone games as efficacious language learning tools.

Figure 6.57 Post-game smartphone efficacy correlation analysis 2

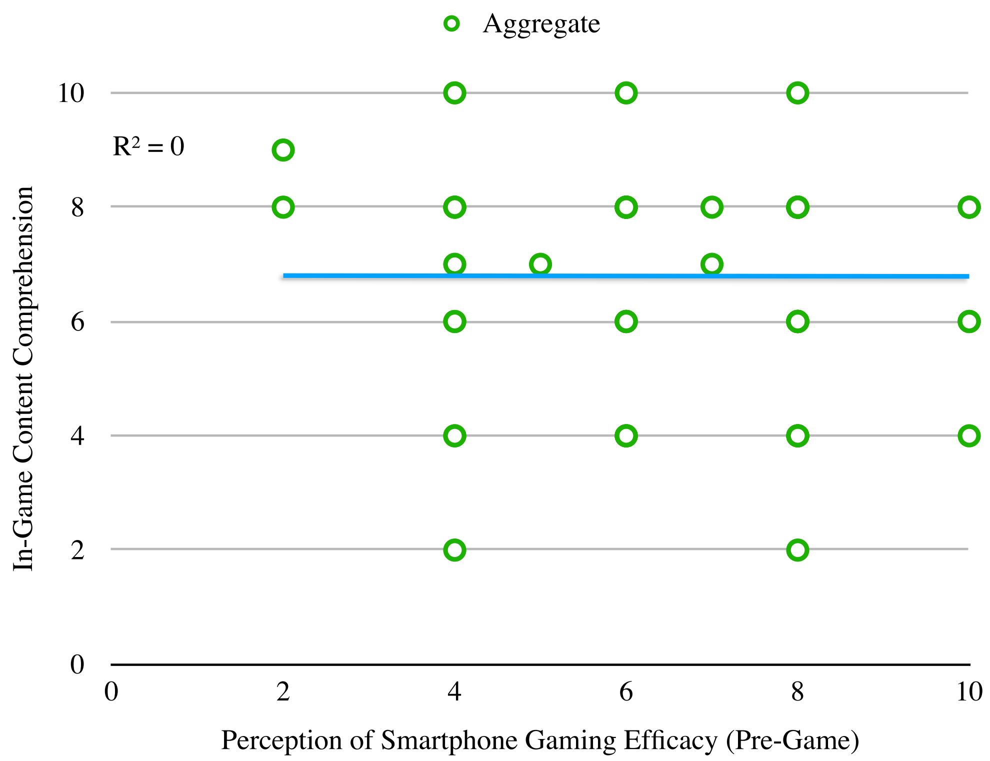


Figure 6.58 Post-game smartphone efficacy correlation analysis 3

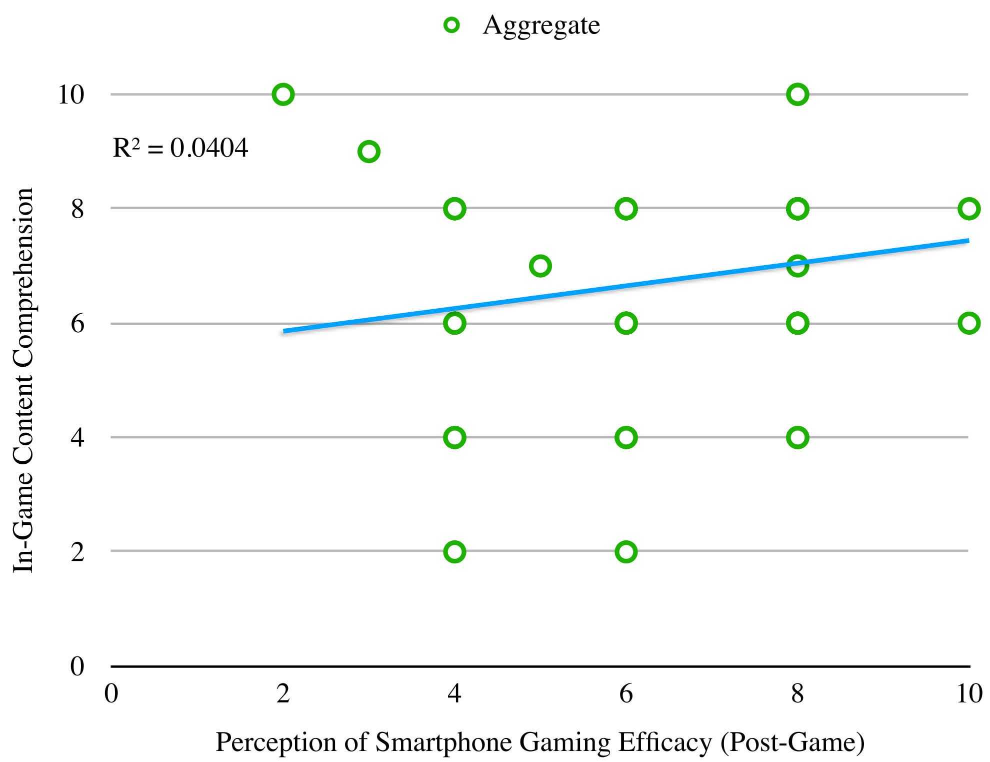


Figure 6.59 Post-game smartphone efficacy correlation analysis 4

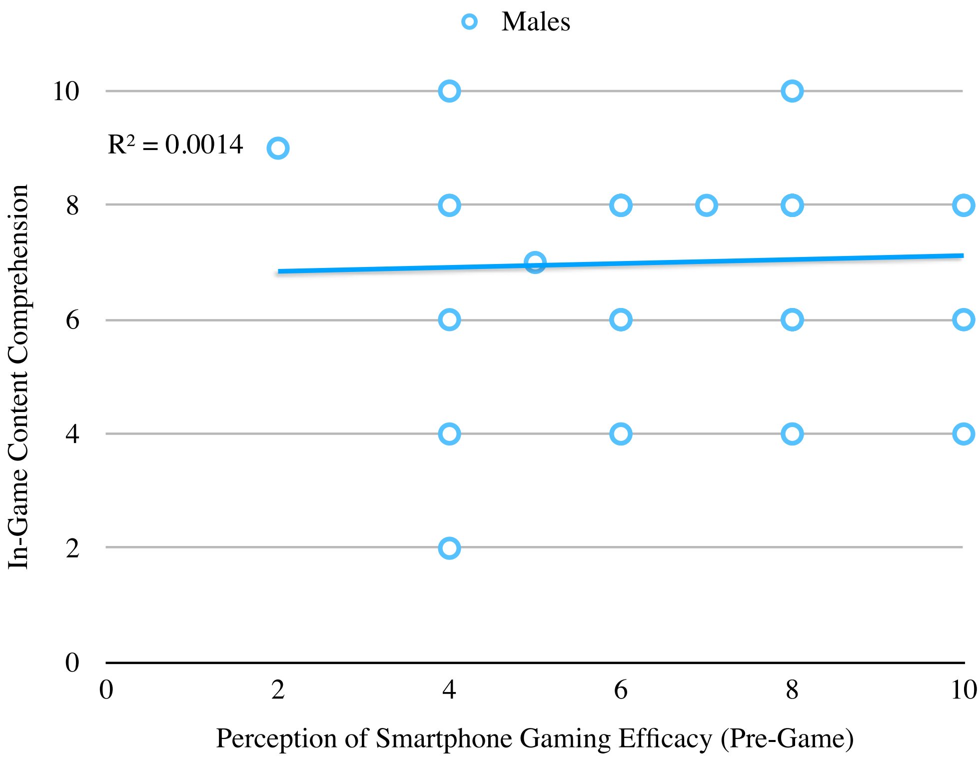


Figure 6.60 Post-game smartphone efficacy correlation analysis 5

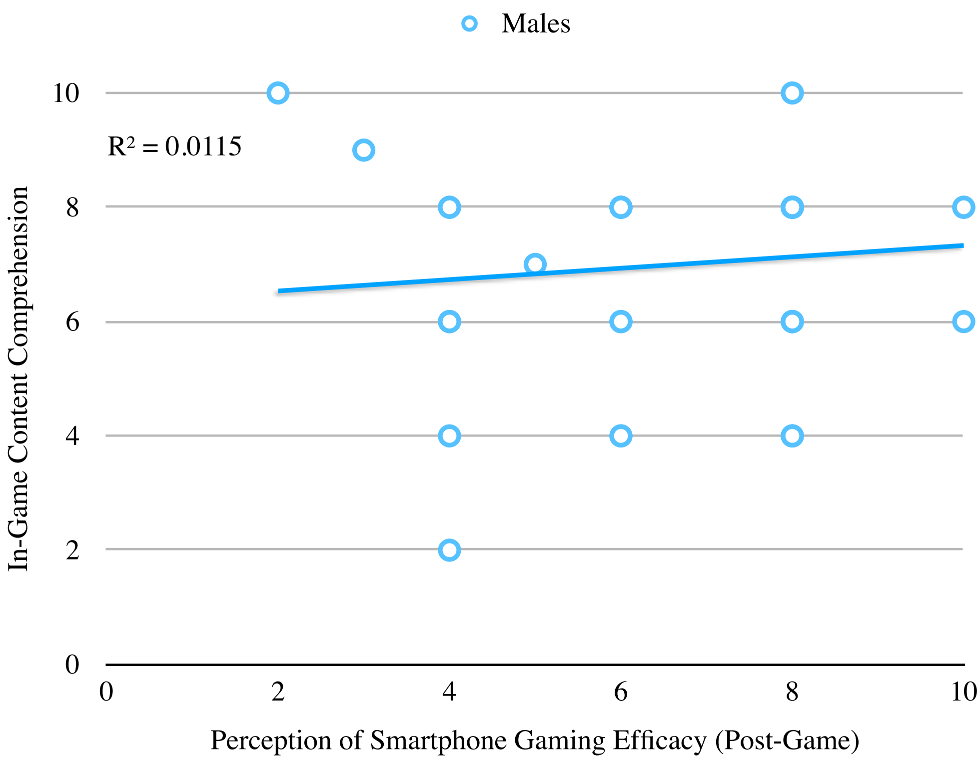


Figure 6.61 Post-game smartphone efficacy correlation analysis 6

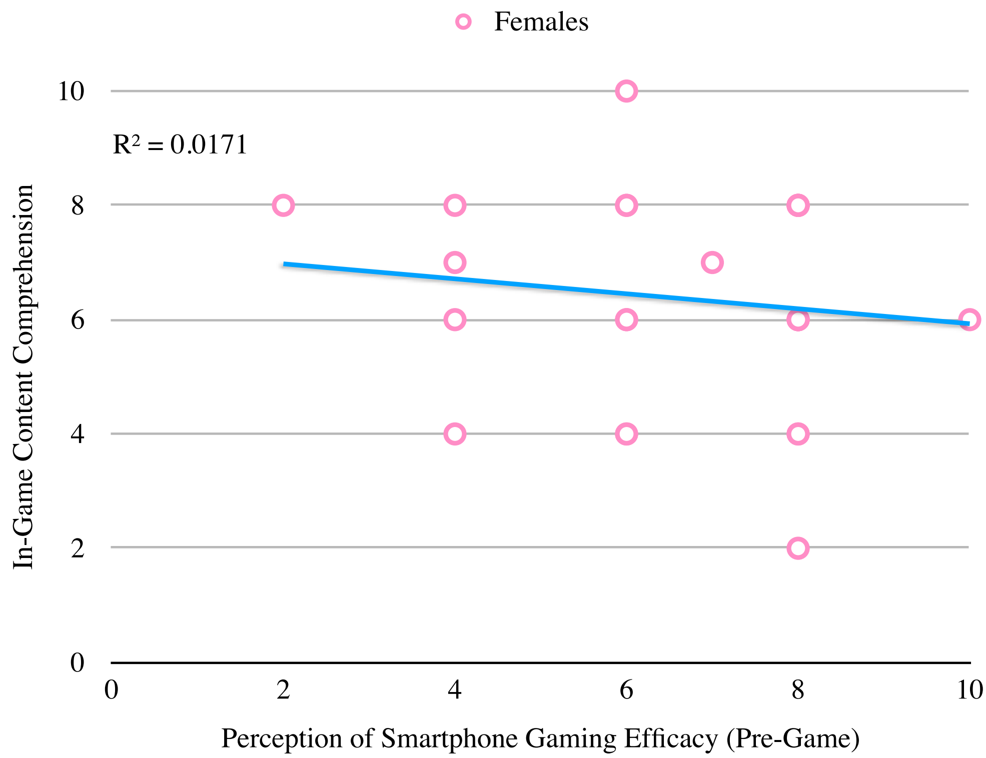
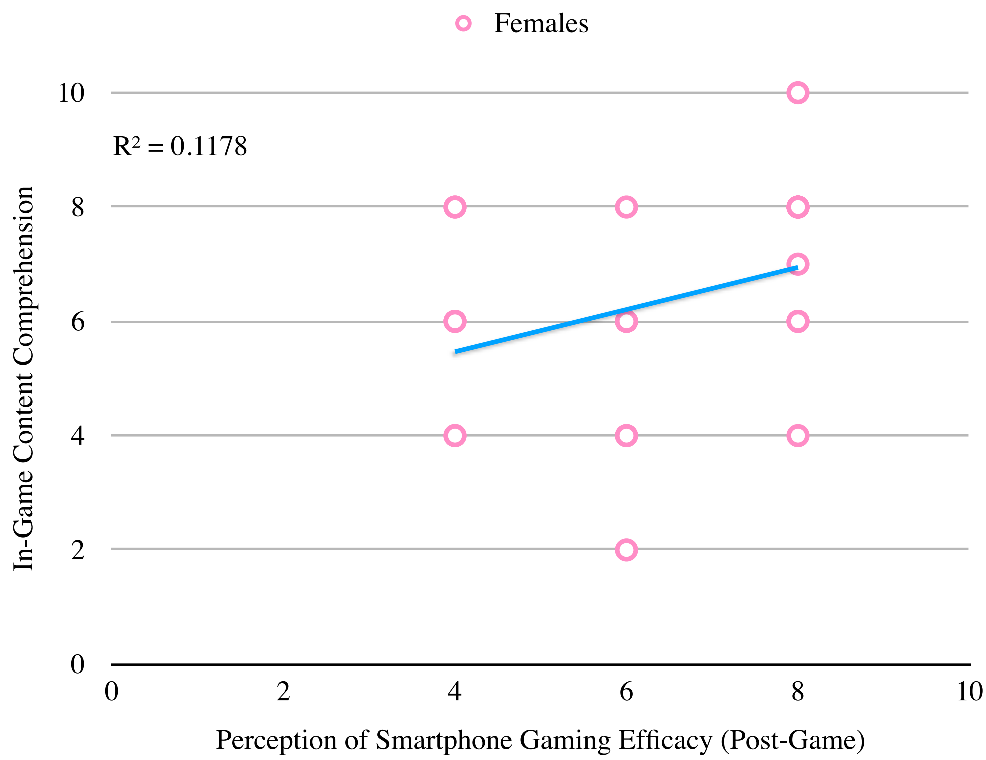


Figure 6.62 Post-game smartphone efficacy correlation analysis 7



A statistically significant, weak positive correlation (r=0.28, p=0.013) was observed between the post-game perception of smartphone gaming efficacy and the perception of difficulty of the in-game English in the aggregate sample (see Figure 6.63). As the correlation between the above-stated values was observed to be negligible and non-significant (r=0.15, p=0.29) in the male sample (see Figure 6.64), the source of this significant positive correlation was found to be largely from the female sample (see Figure 6.65), in which a strong positive correlation was observed (r=0.54, p=0.0038). Calculating the observed z-value between the difference in r-values from both gender groups in the above-stated correlation analysis revealed the difference to be moderately significant (z=1.80, p=0.072). When comparing the above-stated post-game correlations to the pre-game perception of smartphone gaming efficacy values, a significant positive shift in the correlational relationship between the perception of smartphone gaming efficacy and the perception of in-game English difficulty was observed in female sample. In the female sample, the correlation between the pre-game perception of smartphone gaming efficacy values and the perception of the difficulty of in-game English values (r=0.12, p=0.54) was negligible and non-significant (see Figure 6.66). The post-game shift in the female perception of smartphone gaming efficacy relative to pre-game values again suggests that a significant number of females who found the in-game English content to be difficult were also more likely to change their opinion on how efficacious they perceived smartphone gaming to be as a language learning tool. Combined with the significant shifts in the female sample regarding the correlations between the post-game perception of smartphone gaming efficacy and in-game content comprehension (i.e., the more they comprehended, the more efficacious they perceived smartphone gaming to be) adds further detail to the source of the shift in female opinion. The data seem to support an argument suggesting that the females who both found the in-game English to be difficult and comprehended the said content well were more likely to perceive smartphone gaming to be an efficacious language learning tool after having participated in the study program.

Figure 6.63 Post-game smartphone efficacy correlation analysis 8

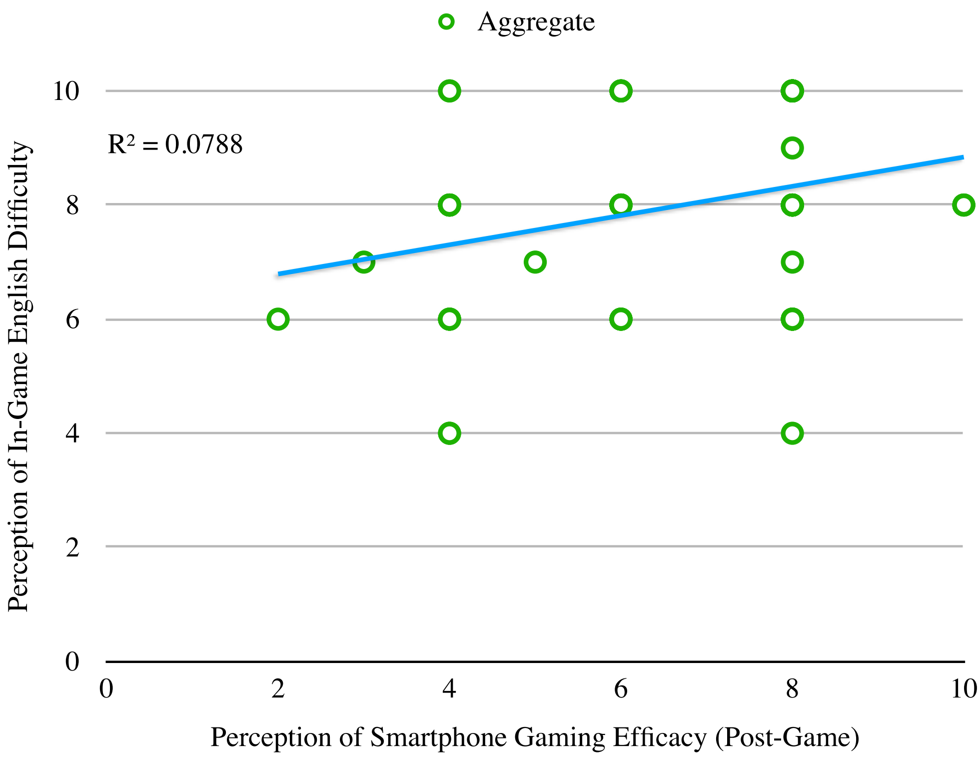


Figure 6.64 Post-game smartphone efficacy correlation analysis 9

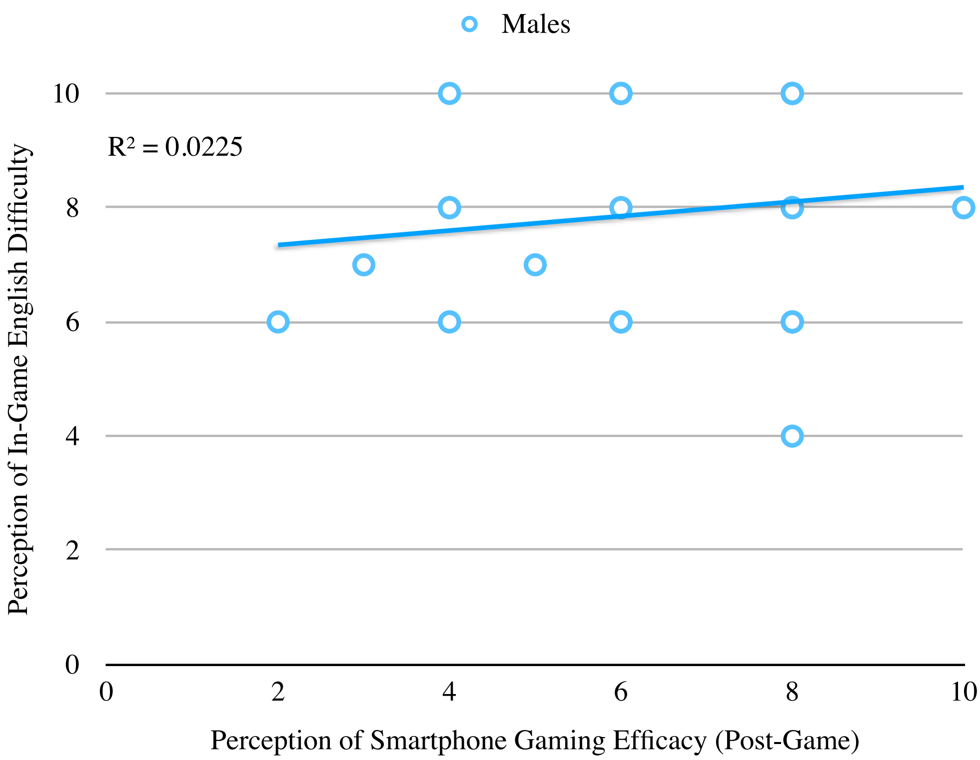


Figure 6.65 Post-game smartphone efficacy correlation analysis 10

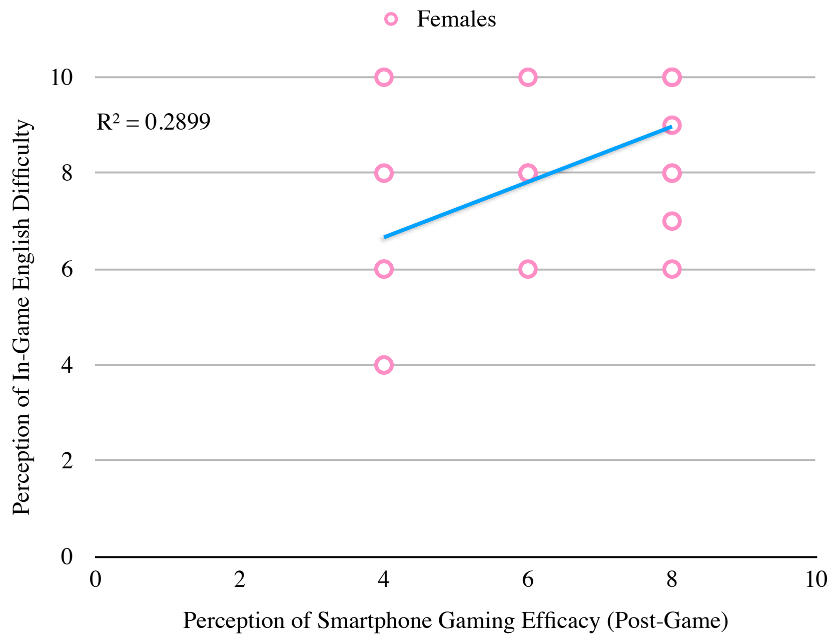
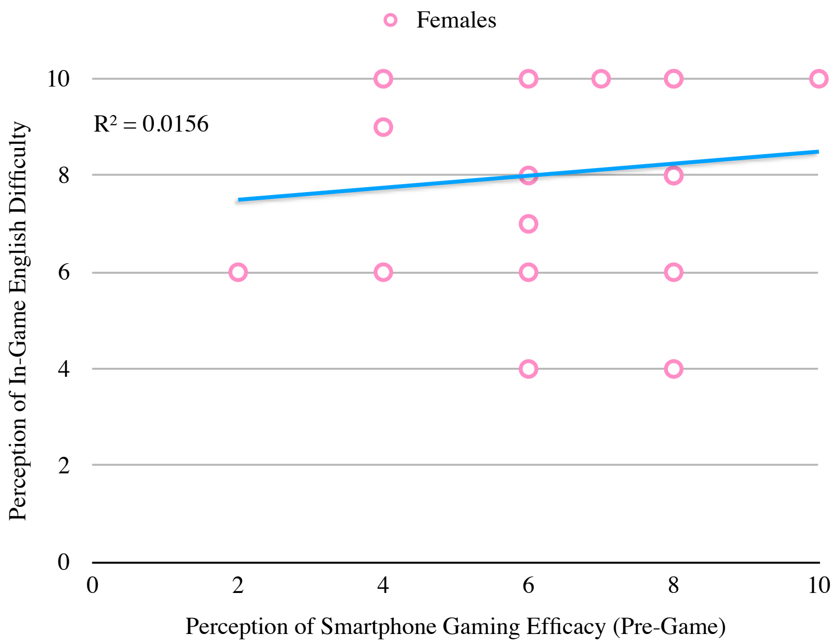


Figure 6.66 Post-game smartphone efficacy correlation analysis 11



Finally, moderate to moderately strong positive correlations between the post-game perception of smartphone gaming efficacy values and the perception of English growth values were observed in the aggregate (r=0.34, p=0.002) (see Figure 6.67) and male samples (r=0.47, p=0.00051) (see Figure 6.68). The above correlation was observed to be negligible and non-significant in the female sample (r=0.12, p=0.55). The source of the statistically significant, positive correlation between the two categories described above is therefore largely from the male sample. When comparing the correlation between the post-game perception of smartphone gaming efficacy values and the perception of English growth values with pre-game values, similar statistically significant, positive correlations were observed in the aggregate (r=0.25, p=0.028) (see Figure 6.26) and male samples (r=0.28, p=0.046) (see Figure 6.27). Similarly, there was a negligible, non-significant correlation for pre-game values in the female sample (r=0.18, p=0.37). Most notably, the strength of the correlation between the post-game perception of smartphone gaming efficacy values and the perception of English growth increased in the male sample. In contrast to the data in the post-game smartphone gaming efficacy analyses which indicate that females' perceptions of the efficacy of smartphone gaming as a language learning tool may in part be affected by both the comprehension and content difficulty of the game, males' perceptions of the efficacy of smartphone gaming were found to be more affected by how much they perceived their English skill to grow as a result of playing the game.

Figure 6.67 Post-game smartphone efficacy correlation analysis 12

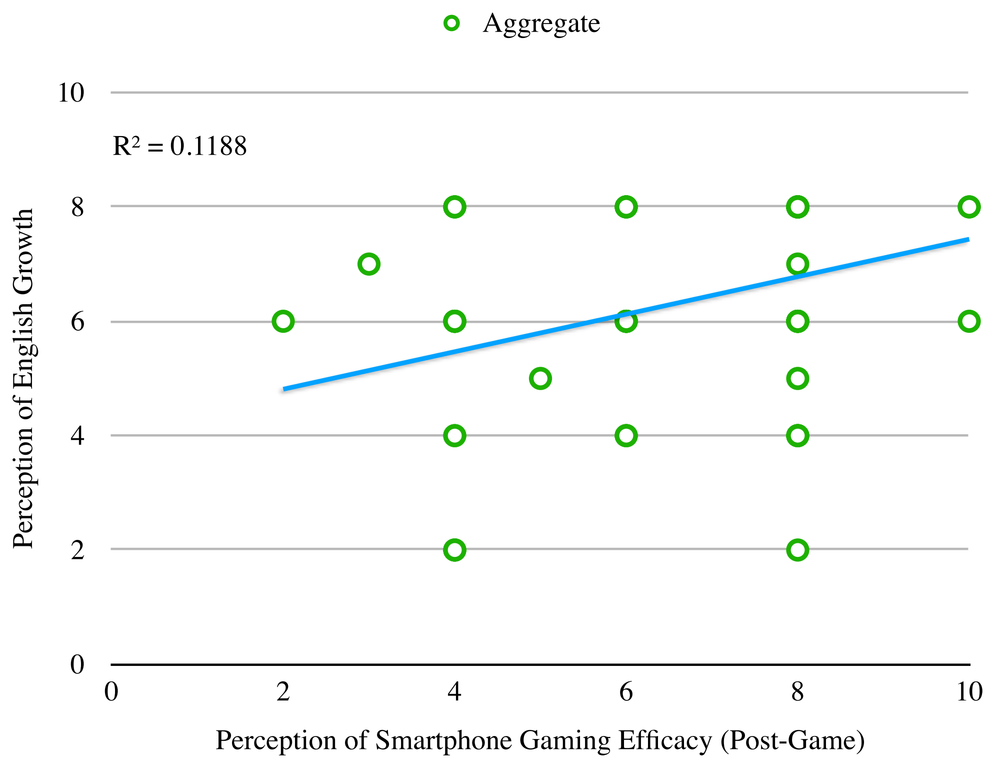
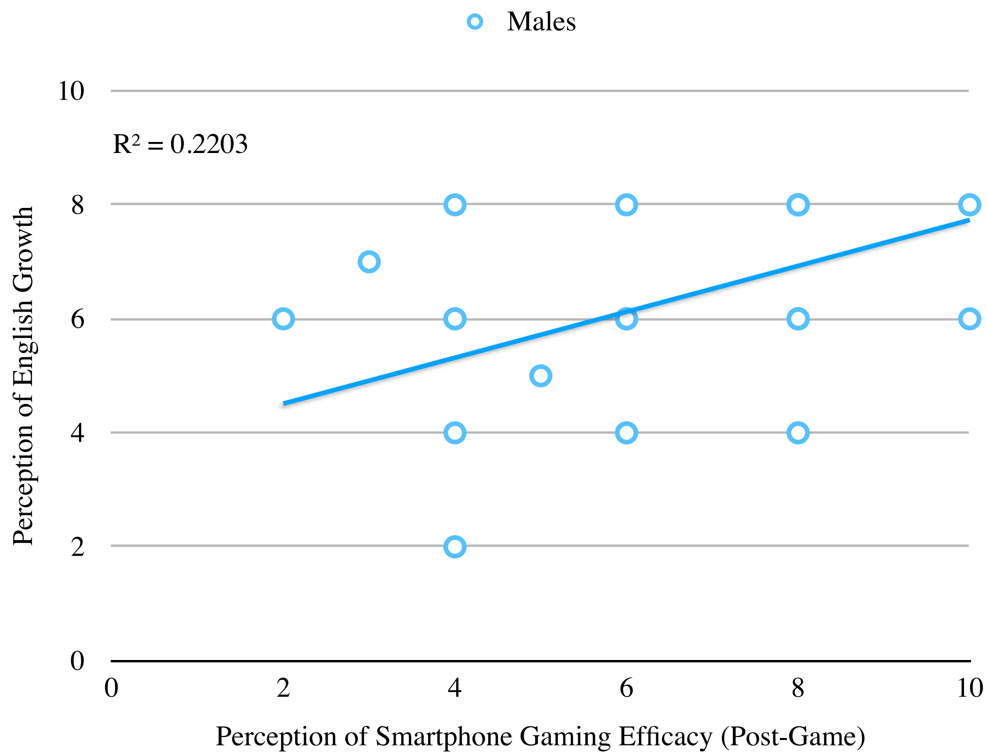


Figure 6.68 Post-game smartphone efficacy correlation analysis 13



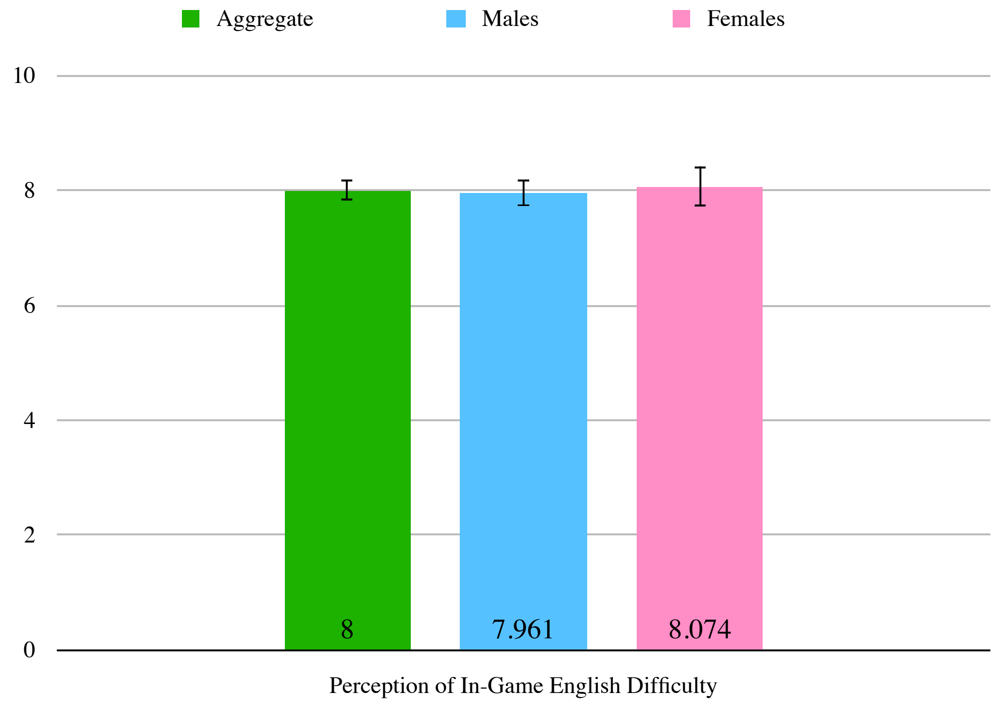
The females' shifts in the perception of smartphone gaming efficacy from pre-game to post-game, although not reflected in the Likert response averages for both pre- and post-game values (see Figure 6.17 and Figure 6.18), are evident in the lack of a statistically significant, positive correlation between pre-game and post-game perception of smartphone gaming efficacy values in the female sample (r=0.08, p=0.69). However, the preservation of approximately identical Likert response averages from pre-game (6.63) to post-game (6.44), means that the sample is populated with roughly equal amounts of females who, after participating in the study program, changed their opinion in positive and negative manners (i.e., some thought smartphone games were more efficacious language learning tools after the study program, others thought them less). For the females in the post-game smartphone efficacy dataset, content comprehension (as is evidenced by the increases in positive correlations in the correlation analyses with comprehension test scores and in-game content comprehension values) and the difficulty of the in-game English (as is evidenced by the increased positive correlation in the correlation analysis with the content difficulty values) were identified as correlational factors.

In contrast to the shift in female opinions regarding the efficacy of smartphone gaming upon completing the study program, the male sample displayed a moderately strong positive correlation (r=0.42, p=0.0022) between pre- and post-game smartphone gaming efficacy values (see Figure 6.25). These data suggest, as would also be supported by the approximately identical Likert response averages from pre-game (7.06) to post-game (6.86), that males who perceived smartphone games to be/not be efficacious language learning tools were likely to continue to believe so upon completing the study. This interpretation of the data would also suggest that the females in the data were more likely to change their opinion regarding smartphone gaming efficacy based on the LIS gaming intervention presented in this study. This is not to say, however, that the male sample was wholly without shifts in smartphone gaming efficacy perception. In particular, the moderate increase in the positive correlation to the perception of English growth values from pre-game (r=0.28, p=0.046) to post-game (r=0.47, p=0.00051) suggests that how much the participants perceived to gain linguistically from playing the game was influential in the post-game perception of smartphone gaming efficacy values.

6.11 Analysis of in-game English difficulty

The following analysis of the students' perceptions of in-game English difficulty was based on data gathered in the post-game survey. Students were asked to rate how difficult they perceived the English used in LIS to be on a 10-point Likert scale (1=not at all difficult 10=very difficult), the results of which are summarized in Figure 6.69. Within-group, two-sample, two-tailed t-tests were used to test for significant differences in Likert values between the gender groups, for which no significant differences were observed.

Figure 6.69 Perception of in-game English difficulty Likert responses



Correlation analyses were also conducted for the perception of in-game English difficulty survey question to analyze potential correlations with other Likert scale survey questions as well as with vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.10. Statistically significant correlations were revealed between the perception of in-game English difficulty values, vocabulary test scores, and comprehension test scores. No other significant differences between the *r*-values in the gender groups were observed. The following section will detail the significant findings.

Table 6.10 Correlation analysis of in-game English difficulty

	Aggregate (n=78)		Males (n=51)		Females (n=27)	
Categories	*r*-value	*p*-value	*r*-value	*p*-value	*r*-value	*p*-value
Vocabulary scores	-0.36	0.001	-0.23	0.10	-0.54	0.0036
Comprehension scores	-0.37	0.016	-0.29	0.15	-0.57	0.027
Perception of English growth	0.023	0.84	0.17	0.24	-0.17	0.40

Moderate to strong negative correlations of statistical significance were observed between the perception of in-game English difficulty values and vocabulary scores in both the aggregate (r=-0.36, p=0.001) (see Figure 6.70) and female samples (r=-0.54, p=0.0036) (see Figure 6.71). Although a similar negative correlation between the above-stated factors was observed in the male sample (r=-0.23, p=0.1), the correlation was weak and statistically non-significant. The negative correlations observed across all three samples suggest that the students who found the English used in LIS to be difficult were less likely to perform well on the vocabulary tests. The strong, statistically significant negative correlation in the female sample may suggest that this effect was more pronounced for the females relative to the males.

Figure 6.70 Perception of in-game English difficulty correlation analysis 1

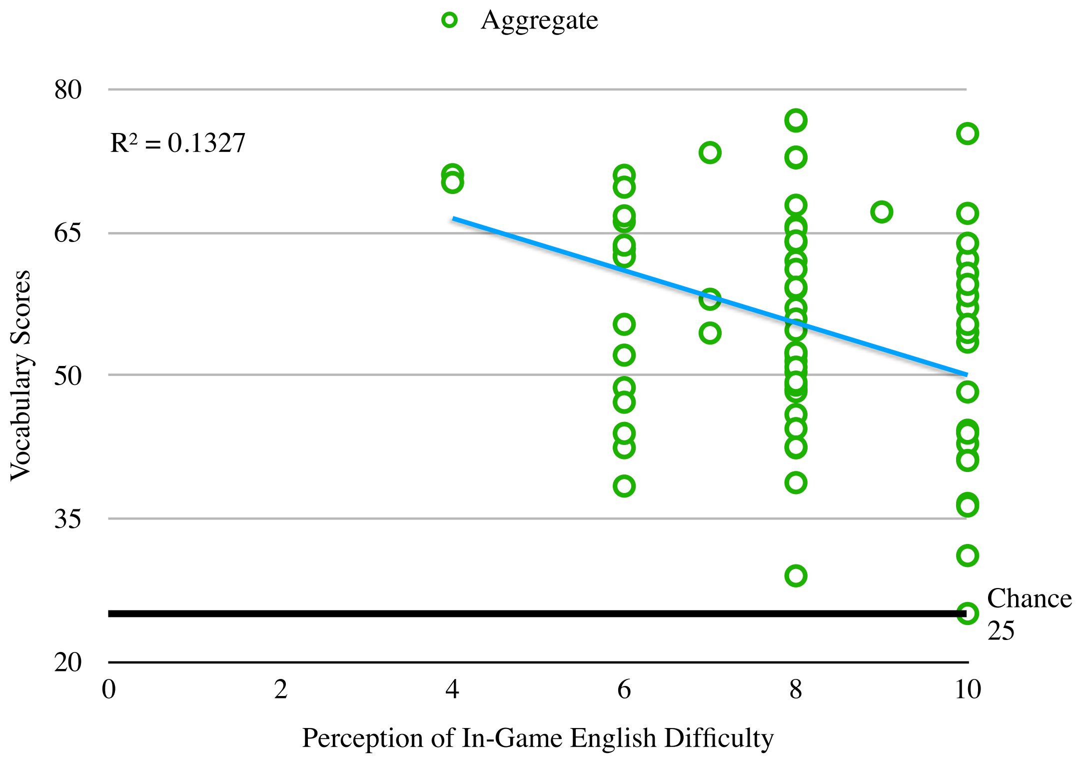
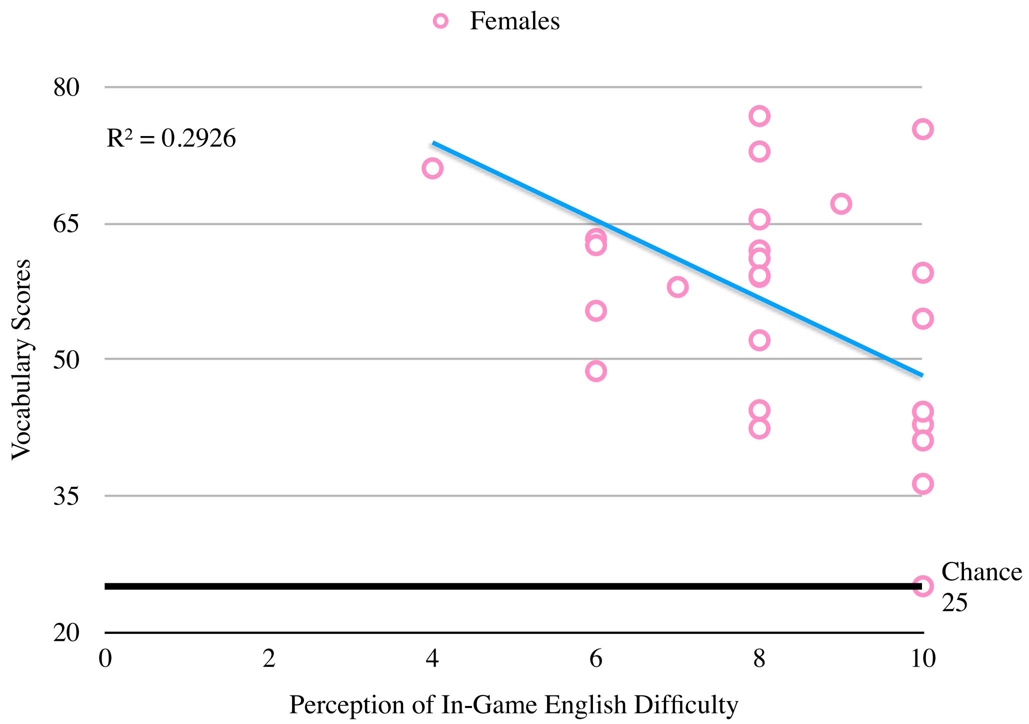


Figure 6.71 Perception of in-game English difficulty correlation analysis 2



Very similar moderate to strong negative correlations of statistical significance were also observed between the perception of in-game English difficulty values and comprehension test scores in both the aggregate (r=-0.37, p=0.016) (see Figure 6.72) and female samples (r=-0.57, p=0.027) (see Figure 6.73). Again, a similar negative correlation between the above-stated factors was observed in the male sample (r=-0.29, p=0.15), but the correlation was weak and statistically non-significant. Much like the negative correlations to the vocabulary test scores, the negative correlations to the comprehension test scores observed across all three samples suggest that students who found the English used in LIS to be difficult were less likely to perform well on the comprehension tests. The strong negative correlation observed in the female sample also suggests that the effect was again more pronounced for the females than the males. These results are fully in line with the moderately strong, statistically significant positive correlations observed between the vocabulary and comprehension test scores (see Figure 5.9) and support the contention that the English proficiency level of the student is correlated to more in-game comprehension and, in turn, better rates of in-game vocabulary acquisition/retention.

Figure 6.72 Perception of in-game English difficulty correlation analysis 3

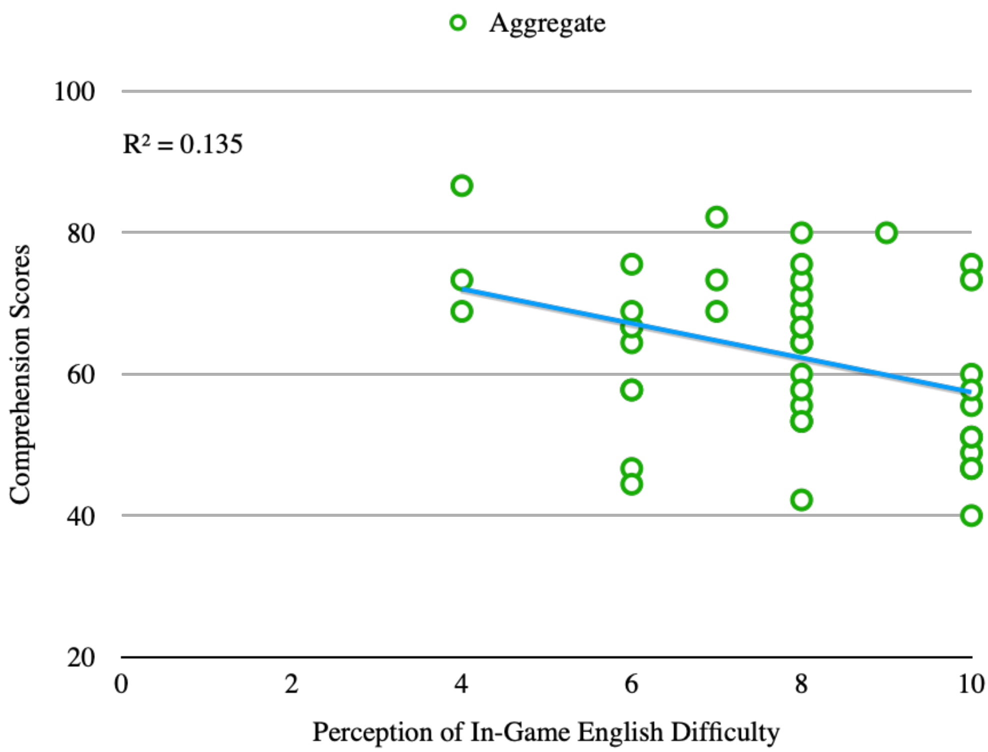
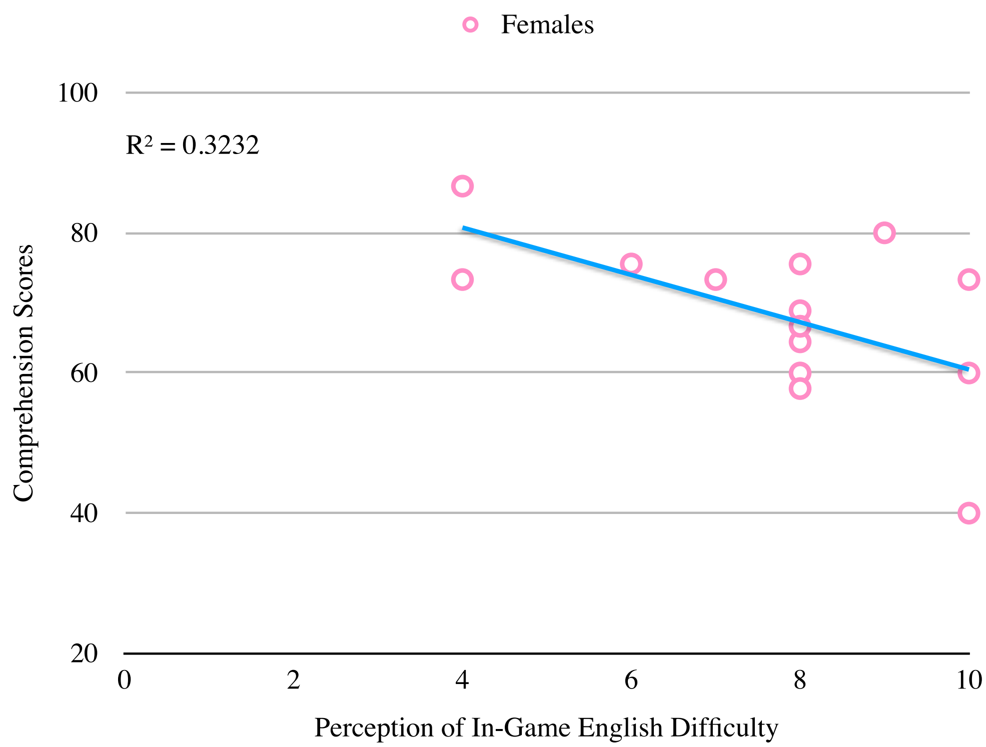


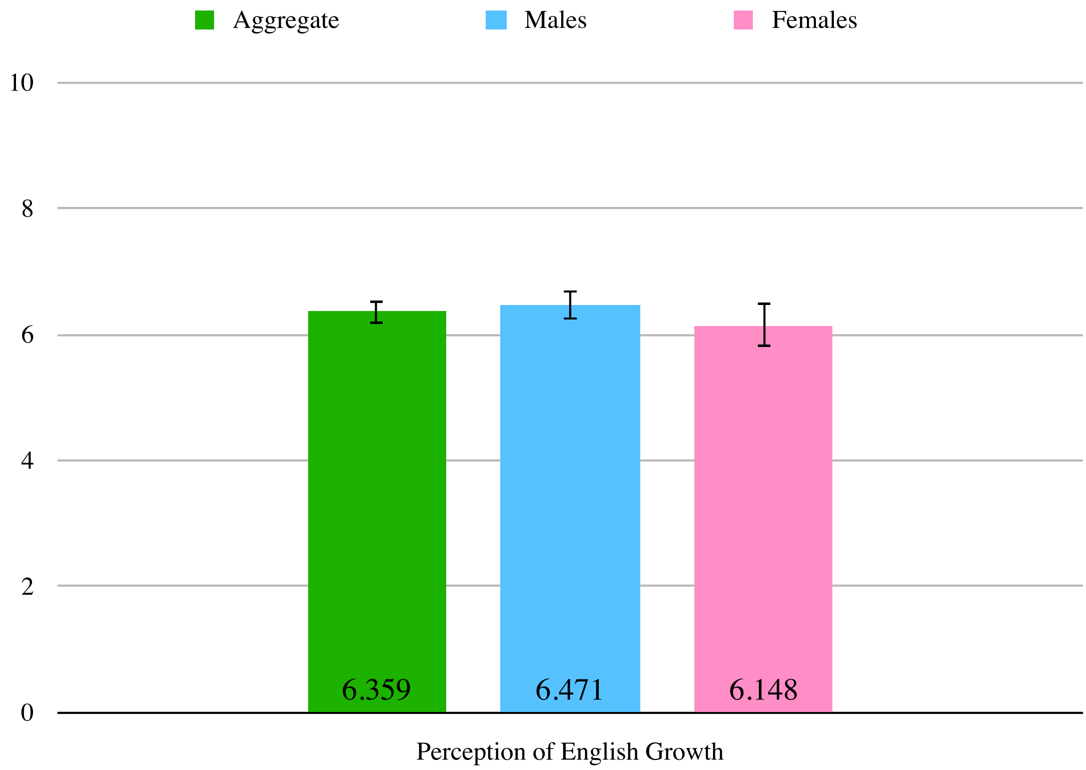
Figure 6.73 Perception of in-game English difficulty correlation analysis 4



6.12 Analysis of perception of English growth from LIS

The following analysis of the students' perceptions of growth in English language proficiency as a result of participating in the present study program was based on data gathered in the post-game survey. Students were asked to rate how much they perceived their English to improve after playing the game on a 10-point Likert scale (1=no growth 10=significant growth), the results of which are summarized in Figure 6.74. Within-group, two-sample, two-tailed *t*-tests were used to test for significant differences in Likert values between the gender groups, for which no significant differences were observed.

Figure 6.74 Perception of English growth Likert responses



Correlation analyses were also conducted for the perception of English growth survey question to analyze potential correlations with the vocabulary and comprehension test scores. The results of these correlation analyses are summarized in Table 6.11. No statistically significant correlations were observed. Despite the non-significant correlations, there were moderately significant differences observed between the *r*-values in the gender groups for vocabulary and comprehension test scores. The following section will detail the moderately significant findings.

Table 6.11 Correlation analysis of perception of English growth from LIS

	Aggregate (*n*=78)		Males (*n*=51)		Females (*n*=27)	
Categories	*r*-value	*p*-value	*r*-value	*p*-value	*r*-value	*p*-value
Vocabulary scores	0.055	0.63	-0.14	0.33	0.31	0.12
Comprehension scores	0.084	0.6	-0.095	0.64	0.32	0.24
Categories with significant gender differences	Observed *z*-value		Fisher's *z* (males)		Fisher's *z* (females)	
Vocabulary scores	1.85		-0.14		0.32	
Comprehension scores	1.73		-0.095		0.34	

There was a negligible, non-significant, negative correlation (r=-0.14, p=0.33) between the perception of in-game English growth values and vocabulary test scores in the male sample (see Figure 6.75). In contrast to this result, there was a weak, non-significant, positive correlation (r=0.31, p=0.12) between the above-stated factors observed in the female sample (see Figure 6.76). Although both of these results are statistically non-significant, the difference between the two r-values was found to be moderately significant (z=1.85, p=0.064). There were also similar correlations observed to the comprehension test scores, with the male sample displaying a negligible, statistically non-significant, negative correlation (r=-0.095, p=0.64) (see Figure 6.77) and the female sample displaying a weak, non-significant, positive correlation (r=0.32, p=0.24) (see Figure 6.78). The differences in the two comprehension test r-values was also found to be moderately significant (z=1.73, p=0.084).

Figure 6.75 Perception of English growth correlation analysis 1

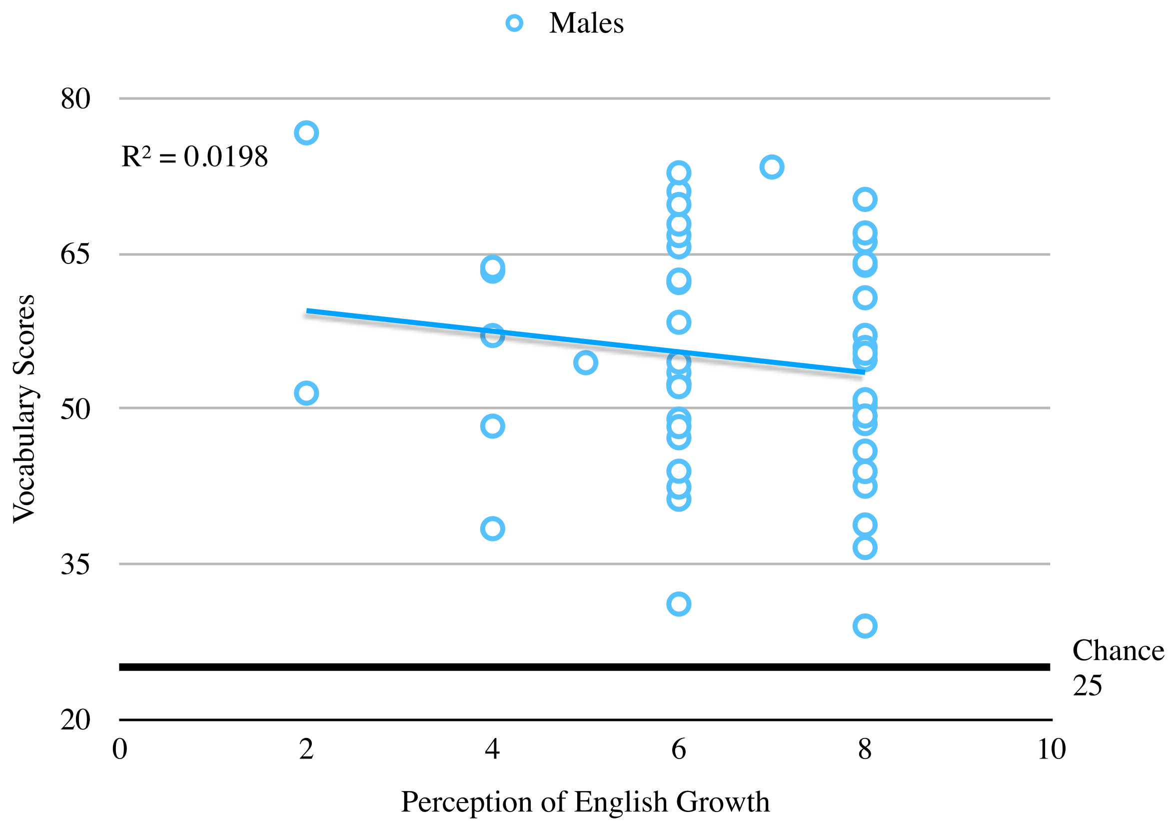


Figure 6.76 Perception of English growth correlation analysis 2

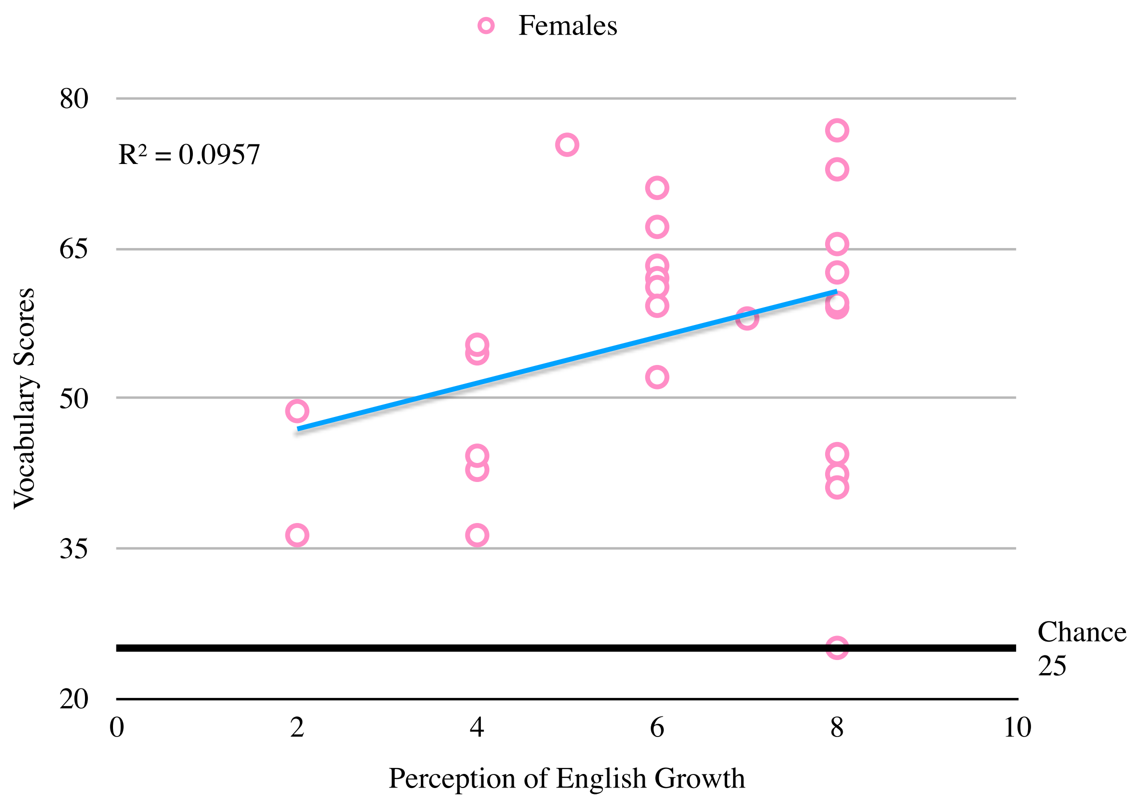


Figure 6.77 Perception of English growth correlation analysis 3

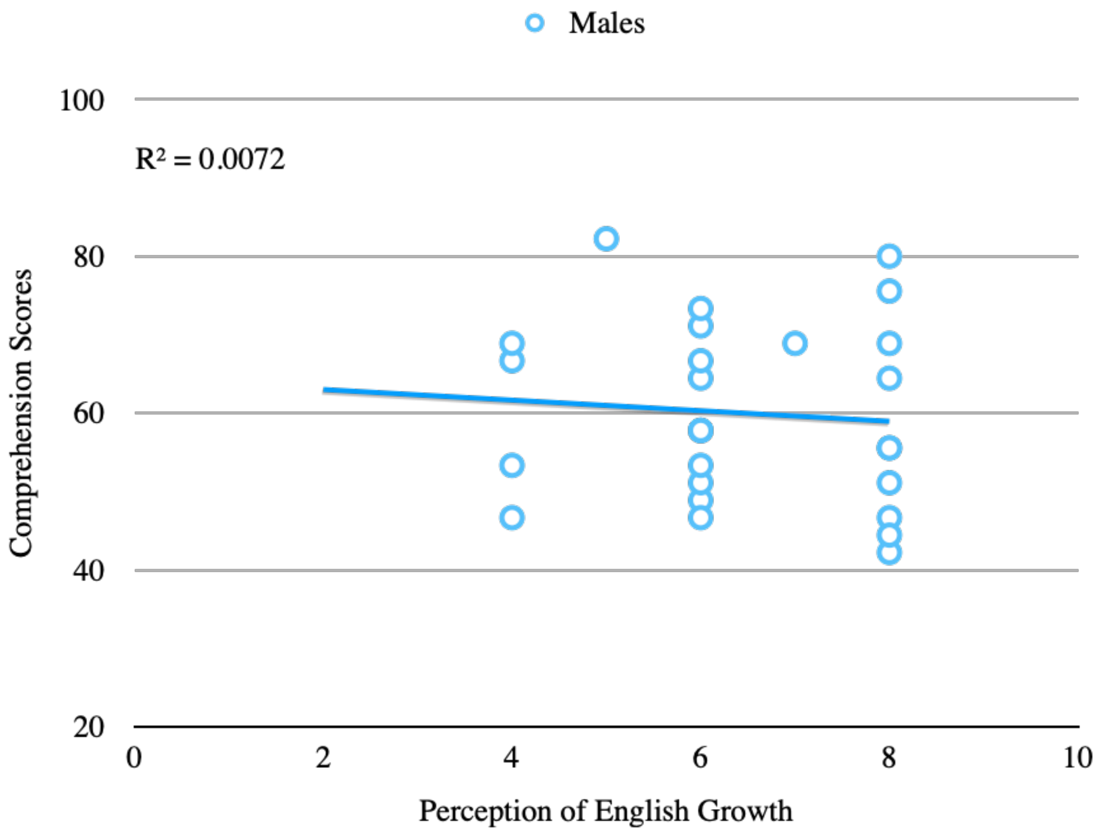


Figure 6.78 Perception of English growth correlation analysis 4

These results suggest a moderately significant difference in how the male and female samples perceived their English growth as a result of the game and their subsequent vocabulary/comprehension test scores. However, the positive correlations observed in the female sample are more pronounced than the negative correlations observed in the male sample. These results also provide limited evidence which may suggest that the females who performed well/poorly on the vocabulary/comprehension tests were slightly more/less likely to perceive that their English improved as a result of the LIS gaming intervention presented in this study, especially relative to the male sample. Overall, these moderately significant gender differences also highlight the slightly more pronounced positive correlations between vocabulary and comprehension test scores observed in the female sample (see Figure 5.11) relative to the male sample (see Figure 5.10). These gender differences may provide limited evidence to argue that the females valued the testing intervention results more than the males in regard to how much they perceived their English proficiency to improve as a result of the LIS gaming intervention presented in this study.

6.13 Summary of perceptions of smartphone gaming

This chapter presented the results of the pre-/post-test surveys to analyze how the participants' perceptions of and attitudes towards smartphone gaming correlated to the LIS gaming intervention. Significant correlations and gender differences from all correlation analyses are summarized in Table 6.12. An in-depth interpretation of these results is provided in the conclusion chapter. Of particular note, however, is the relatively large number of significant gender differences within the perception of English growth category. These gender differences particularly highlight the complexities of designing efficacious DGBL curricula which best suit the needs of everyone in the class.

Table 6.12 Summary of significant correlations and gender differences from pre- and post-test surveys

		1	2	3	4	5	6	7	8	9
1	Pre-Game Smartphone Engagement	x					Pos.	Pos. (M) Neg. (F)		Pos. (M) Neg. (F)
2	Vocabulary Scores		x	Pos.	Neg. (M)			Neg.	Neg. (F)	Neg. (M) Pos. (F)
3	Comprehension Scores		Pos.	x	Neg.	Pos. (M) Neg. (F)			Neg. (F)	Neg. (M) Pos. (F)
4	Pre-Game Perception of Smartphone Gaming Efficacy		Neg. (M)	Neg.	x			Pos. (M)		Pos. (M)
5	Perception of In-Game Content Comprehension			Pos. (M) Neg. (F)		x				Pos. (M)
6	Post-Game Smartphone Engagement	Pos.					x			Pos.
7	Post-Game Perception of Smartphone Gaming Efficacy	Pos. (M) Neg. (F)	Neg.		Pos. (M)			x	Pos. (F)	Pos. (M)
8	In-Game English Content Difficulty		Neg. (F)	Neg. (F)				Pos. (F)	x	
9	Perception of English Growth	Pos. (M) Neg. (F)	Neg. (M) Pos. (F)	Neg. (M) Pos. (F)	Pos. (M)	Pos. (M)	Pos.	Pos. (M)		x

Note. (M) = male sample, (F) = female sample, no parentheses = aggregate sample, pos. = significant positive correlation, neg. = significant negative correlation

The ANOVA analyses revealed significant variations within the LG group sample: gaming location, prior English smartphone gaming experience, and prior LIS gaming experience. Specifically, the data revealed that gaming in familiar locations such as at home was correlated to better performance on the vocabulary test protocol. Furthermore, familiarity with smartphone gaming and prior experience with LIS was correlated to better performance on the vocabulary test protocols, especially for males. The issue of how motivating the game is to play was also shown to be correlated to better performance, especially for males. The implications of these results are further analyzed in the conclusion chapter.

Finally, the LIS gaming intervention was shown to elicit subtle shifts in both opinions of and attitudes towards smartphone gaming. The results showed significant increases in the number of students who thought that EFL listening skills could be improved with smartphone gaming. In the correlation analyses, results showed a negative correlation between participants' pre-game level of smartphone engagement and performance on the vocabulary and comprehension tests, whereas positive correlations were observed between participants' post-game intention to engage with smartphone gaming and performance on the vocabulary and comprehension tests. Furthermore, there were increases in the strength of positive correlations between pre- to post-game perceptions of smartphone gaming efficacy and the perceptions of in-game comprehension/content difficulty (especially for females). These results, which will be analyzed further in the conclusion chapter, lend credence to the interpretation that the LIS gaming intervention positively affected students' perceptions of and attitudes towards smartphone gaming as a language learning intervention.

Chapter 7 Results for perceptions of US culture

The following chapter will detail the results of the data analysis which directly address RQ4: How will the depictions of US culture in the narrative content of LIS affect how the students inform their impressions on US culture? Specifically, the following sections will present analyses of qualitative data gathered from the pre- and post-test surveys from the LG group to address the research question above.

7.1 Analysis of perception of US culture

In order to measure how the participants' perceptions of US culture were influenced by playing LIS in the present study program, pre-game survey questions asked students to list their impressions on US culture, the US educational system, US students, and what media sources they used to inform said impressions. These pre-game survey questions were followed up with identical post-game survey questions. The survey responses were categorized with recursive analysis, whereby categories of responses identified in one response were then applied to all responses to quantify said categories for the entire dataset. These quantified response categories were then compared from pre- to post-game with paired, two-tailed *t*-tests to identify statistically significant shifts in perceptions. The percentage breakdown for gender groups for each category were also analyzed to identify potentially gender-specific responses and significant shifts in the gender breakdown for each response category. The following sections will detail the comparison analyses for each pre-/post-game survey item described above.

For the pre-game survey question asking for the participants' impressions on US culture, 12 response categories were identified and are summarized in Figure 7.1. The gender breakdown for each response category is summarized in Figure 7.2. Compared to the gender percentage for the total dataset (males=65.39%, females=34.62%), there were a number of categories with significant gender-group skews. Response categories with higher male to female ratios included: 2) climate-related habits (e.g., not using umbrellas) (males=100%), 5) health care-related habits (e.g., not wearing masks) (males=100%), and 12) food culture (e.g., large sizes of fast food items) (males=86.67%). Response categories with higher female to male ratios included: 3) social freedom (e.g., freedom of speech) (females=62.5%), no impressions (e.g., I don't know) (females=75%), and media and technology (e.g., movie graphics) (females=50%).

Figure 7.1 Perception of US culture category breakdown (pre-game)

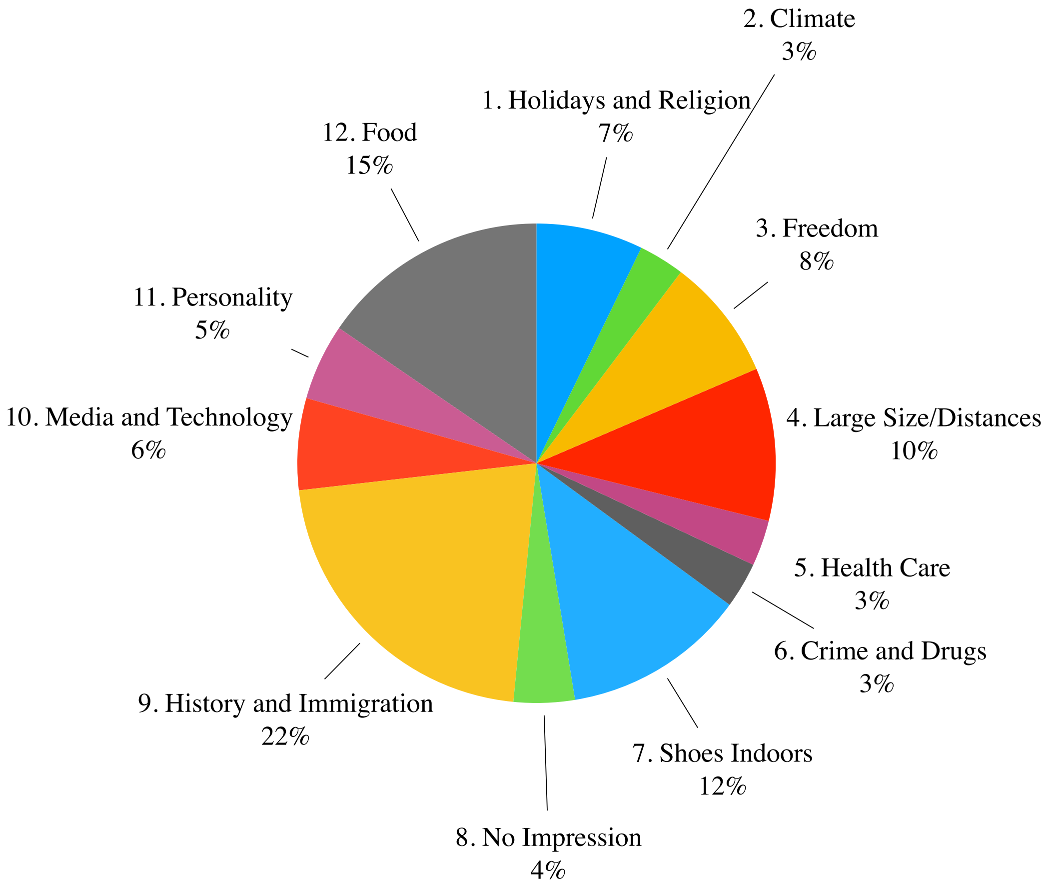
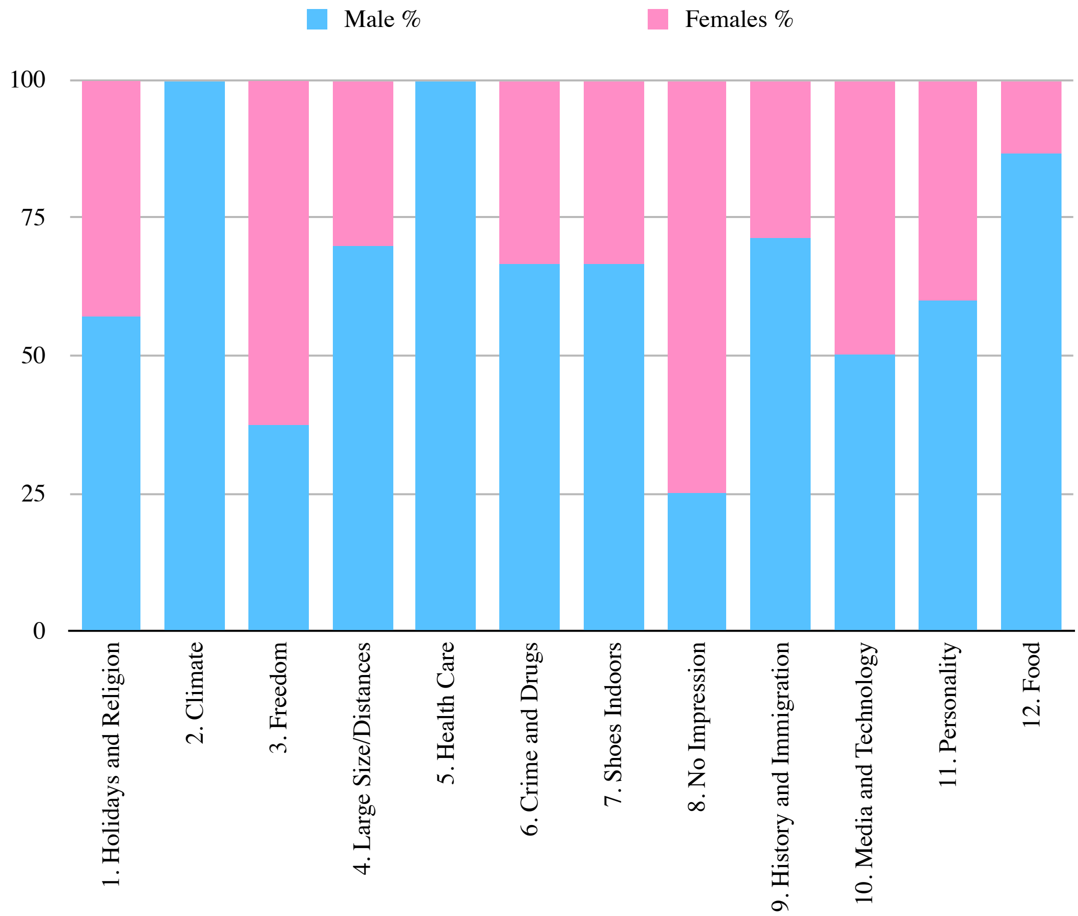


Figure 7.2 Perception of US culture gender breakdown (pre-game)



The post-game survey response category breakdown for the perception of US culture is summarized in Figure 7.3 and the gender breakdown in Figure 7.4. Paired, two-tailed *t*-tests revealed a significant growth (p=0.0059) in the percentage of responses which included mentions of crimes and drugs (category number six) from a pre-game value of 3% to a post-game value of 13%. In particular, the number of post-game responses mentioning guns and marijuana use increased significantly relative to pre-game values. This shift in perception may in part be explained by the in-game narrative content of LIS, which depicts both gun violence within a US school setting and open marijuana use. No other significant changes in the perception of US culture responses were observed. There were, however, notable shifts in the gender spread in response categories 10 and 11. The media and technology response category shrank by approximately 5% in the post-game responses to a total sample of one male whose response included US music. In contrast to the shrinking of response category

10, response category 11 remained approximately the same. However, category 11, the post-game personality response category (e.g., ladies-first mentality), saw a rise in male responses from 60% to 80%. While the significant shift in responses mentioning gun violence and marijuana use was arguably influenced by the in-game narrative content in LIS, it is beyond the scope of the present analysis to determine whether the gender shifts observed in response categories 10 and 11 were influenced by gameplay in any way.

Figure 7.3 Perception of US culture category breakdown (post-game)

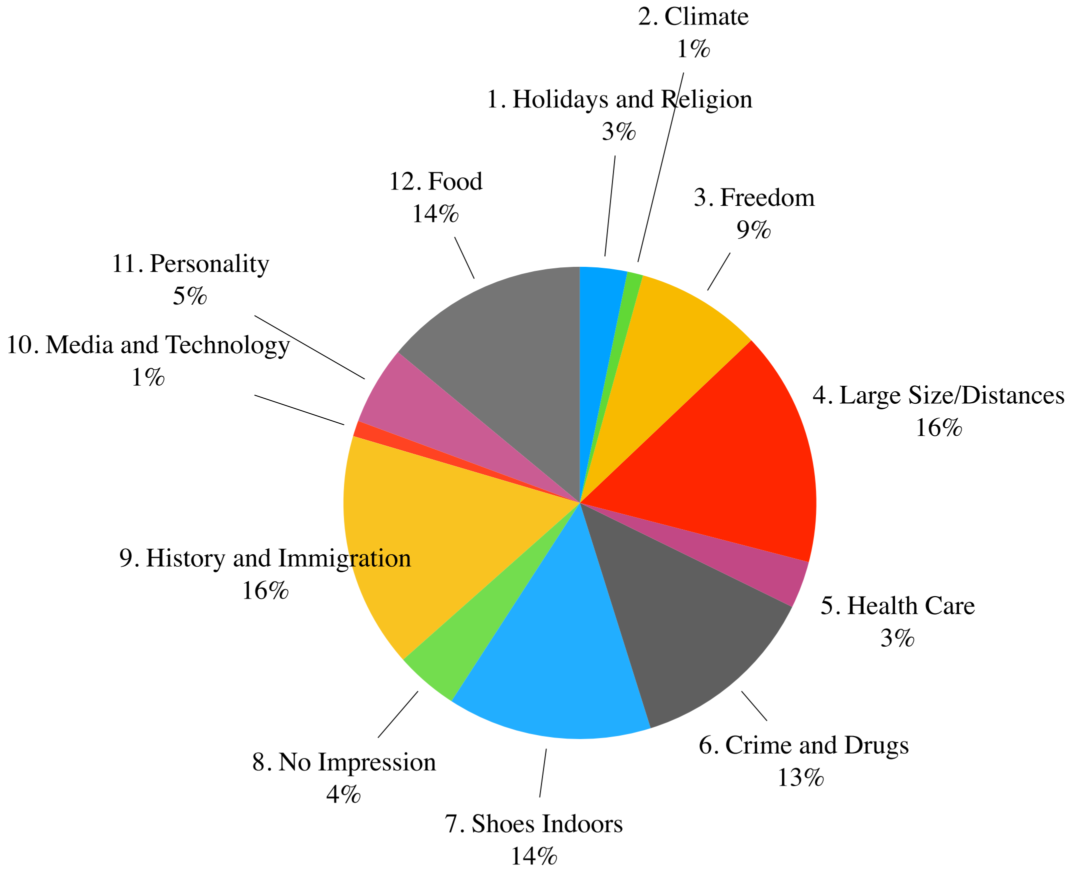
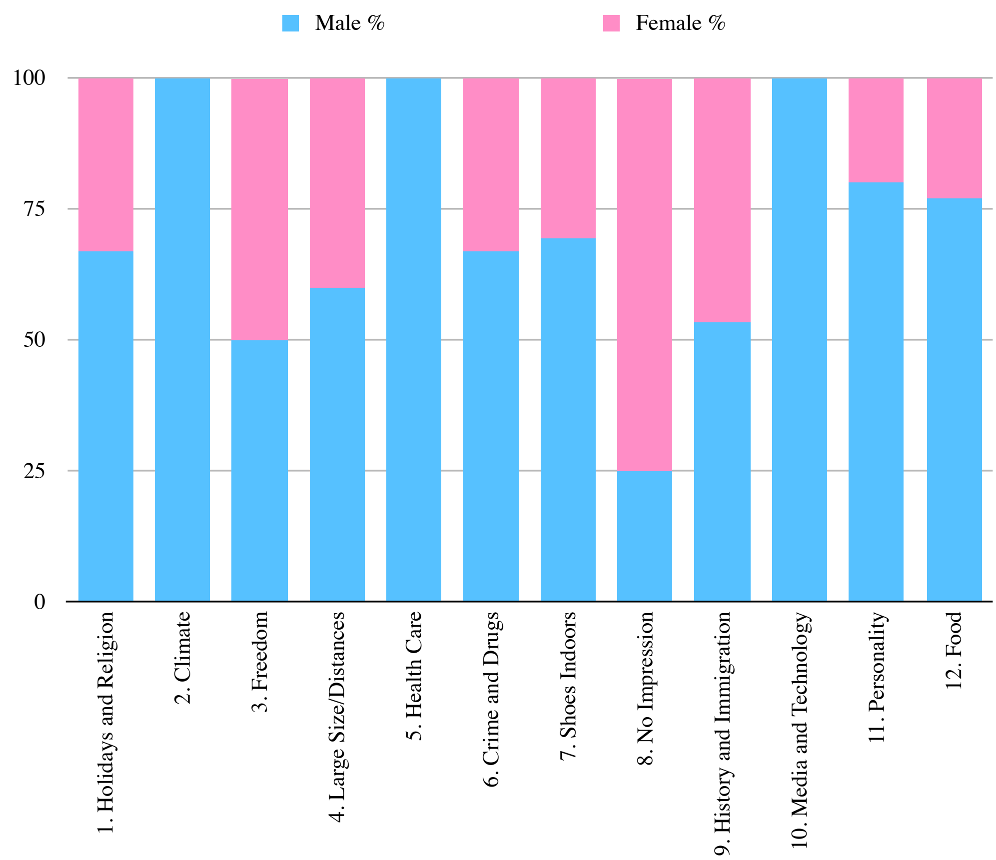


Figure 7.4 Perception of US culture gender breakdown (post-game)



7.2 Analysis of perception of the US educational system

For the pre-game survey question asking for the participants' impressions on the US educational system, eight response categories were identified and are summarized in Figure 7.5. The gender breakdown for each response category is summarized in Figure 7.6. Compared to the gender percentage for the total dataset (males=65.39%, females=34.62%), there were a number of categories with significant gender-group skews. Response categories with higher male to female ratios included: 2) curriculum structure (e.g., US universities are difficult to graduate from) (males=78.57%), 3) communicative educational approach (e.g., makes use of lots of oral presentations) (males=75%), 4) focus on individuals rather than the collective (e.g., no school rules and students are free to take what classes they want) (males=100%), and 7) high cost and importance of school quality (e.g., expensive university tuition and the reputation of the university matters) (males=80%). Response categories with higher female to male ratios included the no impression response category (e.g., I don't know) (females=66.67%).

Figure 7.5 Perception of US educational system category breakdown (pre-game)

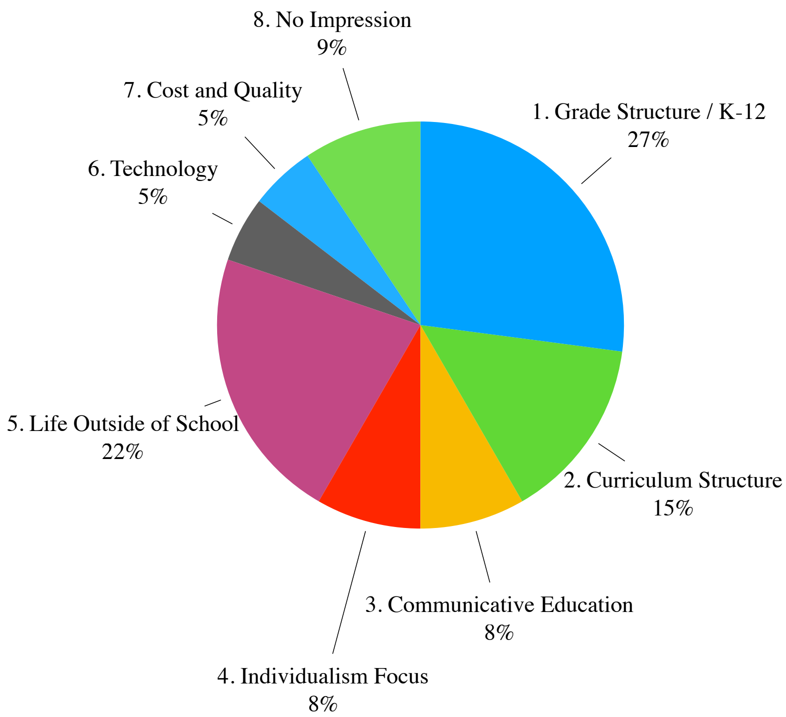
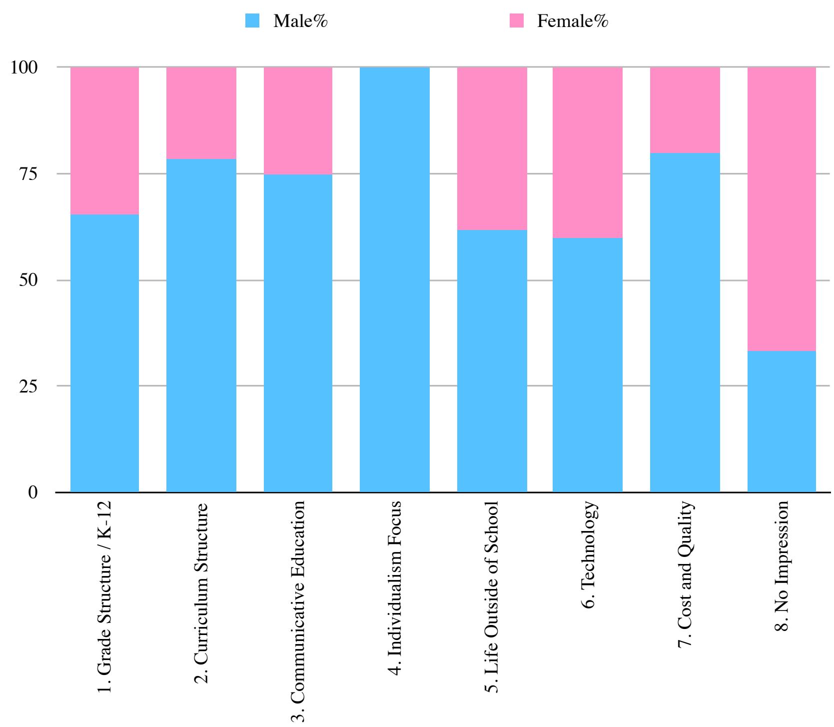


Figure 7.6 Perception of US educational system gender breakdown (pre-game)



The post-game survey response category breakdown for the perception of the US educational system is summarized in Figure 7.7 and the gender breakdown in Figure 7.8. Paired, two-tailed *t*-tests revealed a significant decrease (p=0.034) in the percentage of responses which included mentions of the K-12 grade structure (category number one) from a pre-game value of 27% to a post-game value of 22%. In particular, the percentage of post-game responses mentioning the possibility to skip grades in the US educational system remained approximately the same relative to pre-game survey values, but the total number of responses in that category dropped significantly.

Figure 7.7 Perception of US educational system category breakdown (post-game)

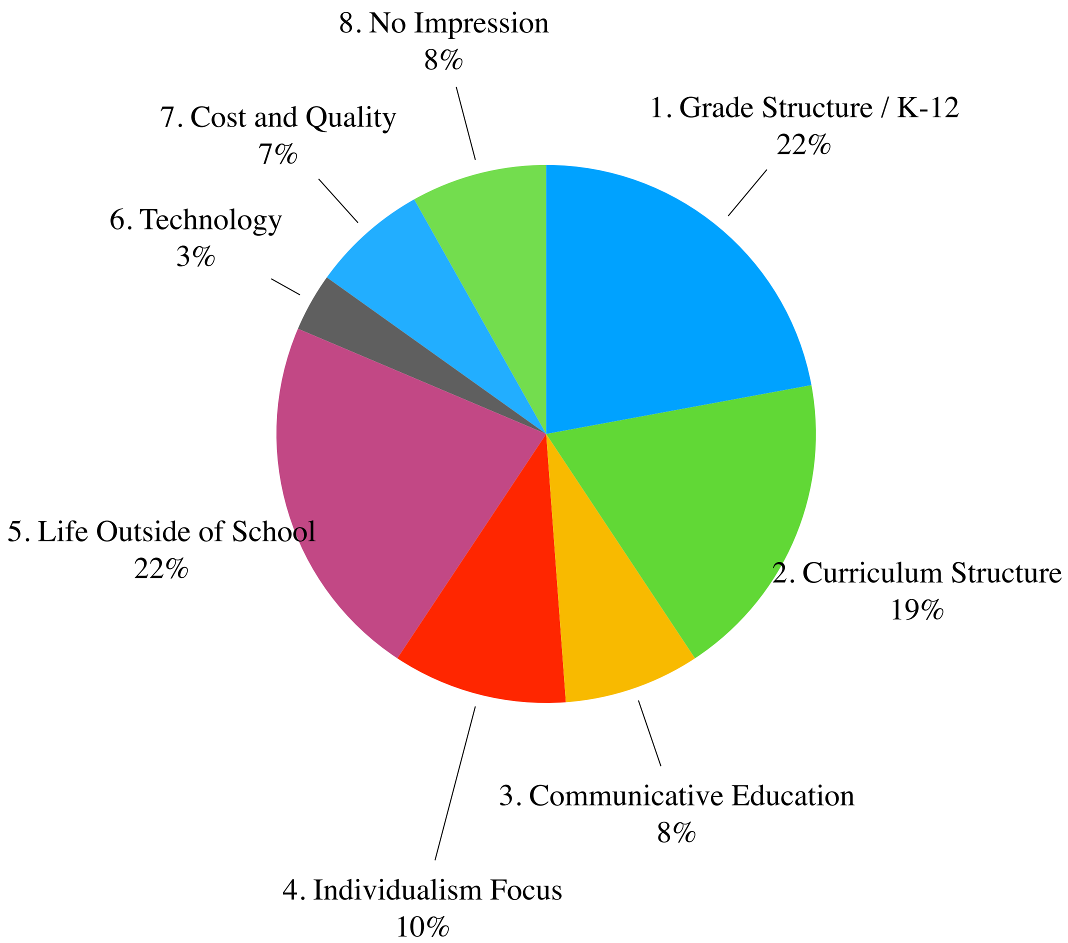
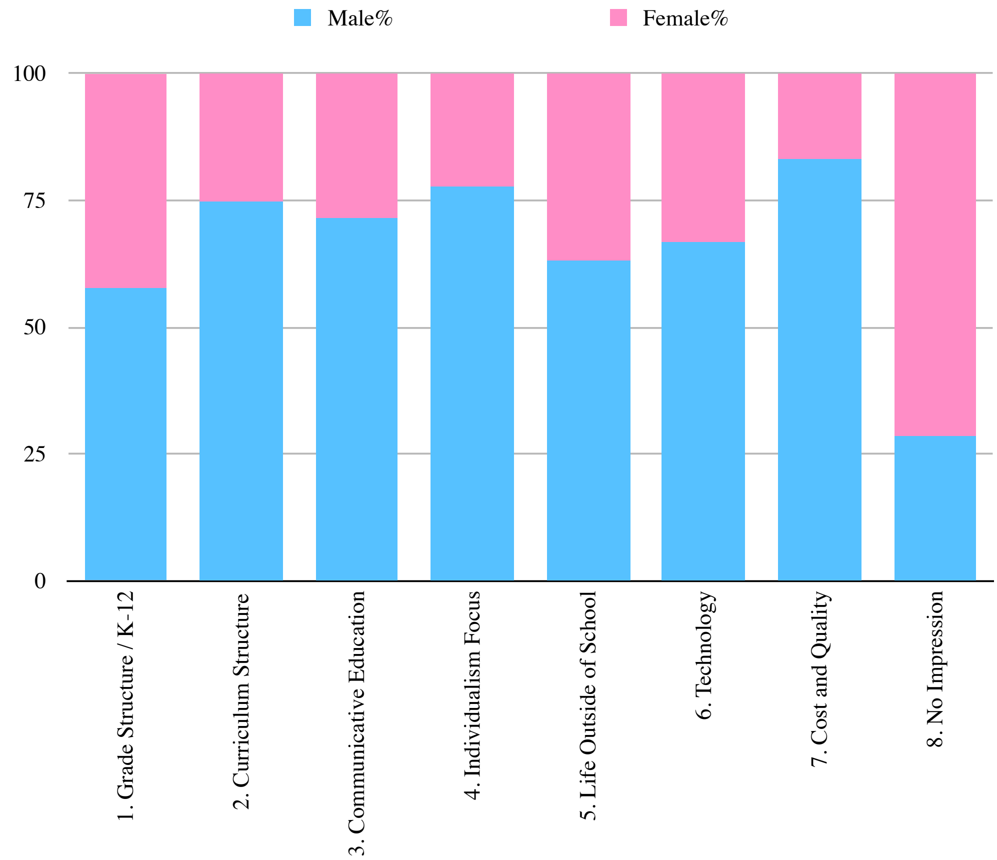


Figure 7.8 Perception of US educational system gender breakdown (post-game)



When considering the in-game narrative content of LIS, the depictions of the US educational system are limited to a high school setting. It may therefore be argued that the lack of reinforcement of the concept of skipping grades within the narrative content in LIS limited such responses. No other significant changes in the perception of the US educational system responses were observed. There was, however, a notable shift in the gender spread in response category four. The post-game individualism focus response category saw a rise in female responses from 0% to 22.22%. In particular, there was a notable increase in responses mentioning the lack of dress code rules (e.g., students are allowed to dye their hair and wear what they want). Increases in these responses may have been influenced by the visual depiction of US high school students in the context of the game. More specifically, the US students depicted in LIS are not wearing uniforms and certain characters have noticeably dyed hair.

7.3 Analysis of perception of US students

For the pre-game survey question asking for the participants' impressions on US students, 10 response categories were identified and are summarized in Figure 7.9. The gender breakdown for each response category is summarized in Figure 7.10. Compared to the gender percentage for the total dataset (males=65.39%, females=34.62%), there were a number of categories with significant gender-group skews. Response categories with higher male to female ratios included: 1) US students' physical stature is large/tall (e.g., they look mature/they are tall) (males=100%), 5) negatively valenced personality (e.g., US football captains are bullies) (males=100%), 6) food consumption (e.g., they eat hamburgers, pizza, and drink cola) (males=100%), 7) fashion/style (e.g., they wear what they want and can dye their hair) (males=85.71%), and 9) cars (e.g., they ride around in cars) (males=100%). Response categories with higher female to male ratios included: 2) strong-willed (e.g., they pursue what they are interested in) (females=52.38%).

Figure 7.9 Perception of US students category breakdown (pre-game)

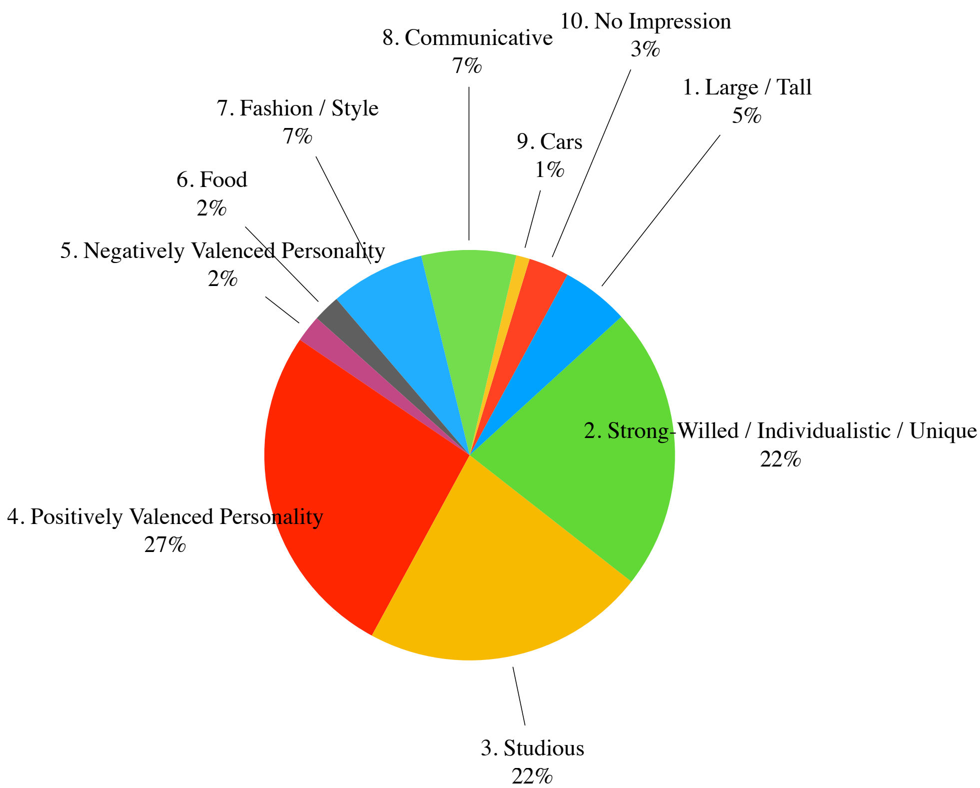
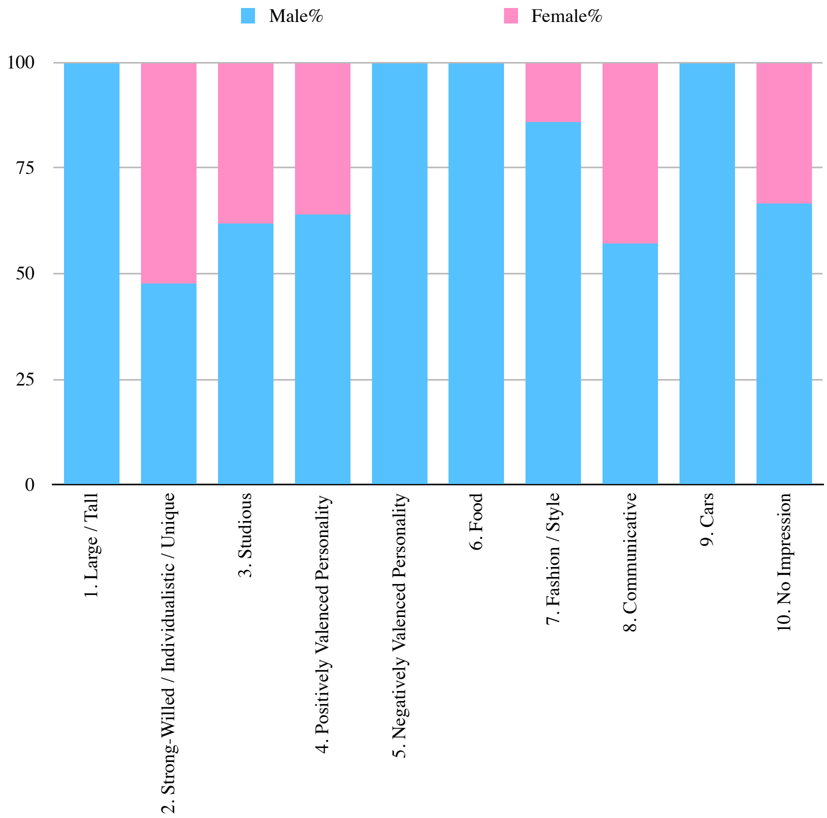


Figure 7.10 Perception of US students gender breakdown (pre-game)



The post-game survey response category breakdown for the perception of US students is summarized in Figure 7.11 and the gender breakdown in Figure 7.12. Paired, two-tailed *t*-tests revealed a significant increase (p=0.017) in the percentage of responses which included mentions of US students being communicative (category number eight) from a pre-game value of 7.45% to a post-game value of 18.09%. In particular, there was a notable increase in the number of responses including how US students actively participate in in-class discussions and presentations. This significant increase in the perception that US students are communicative in class may have been influenced by the narrative content in LIS, specifically by the scenes depicting the photography lesson in Mr. Jefferson's class. Not only does Mr. Jefferson use a Socratic method of education, whereby he continually questions his students to create a dialogue about the lecture theme, but the player must also actively participate in the lesson, effectively allowing the player to virtually experience a common US pedagogical method. No other significant changes in the perception of US students responses were observed.

Figure 7.11 Perception of US students category breakdown (post-game)

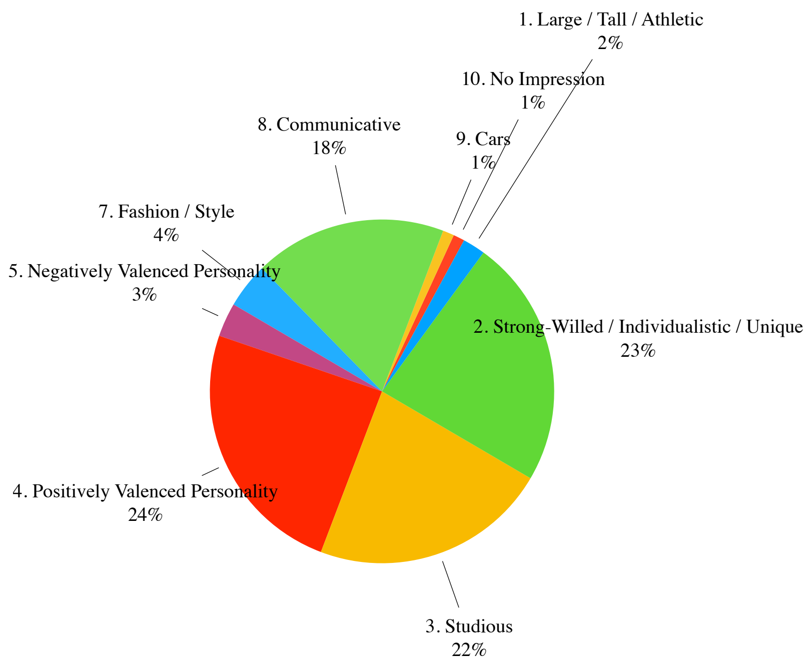
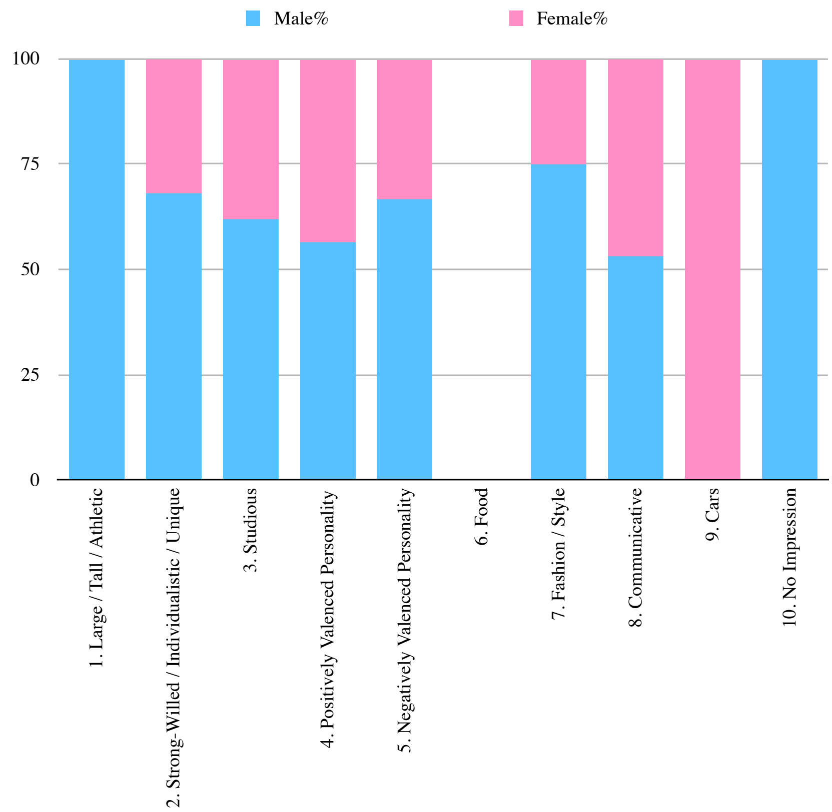


Figure 7.12 Perception of US students gender breakdown (post-game)



There was a notable shift in the gender spread in response category two (strong-willed category). The post-game strong-willed response category saw a rise in male responses from 47.62% to 68.18%, however there were no notable shifts in the types of responses, the most common response in both pre- and post-game surveys being, "They express their own opinions openly." Since the type of responses in the strong-willed response category stayed largely the same pre- to post-game, the narrative content in LIS is an unlikely source of this gender shift. Similarly, the food response category dropped from 2% in the pre-game survey to 0% in the post-game, however since the original population size of the food response category was low to begin with, the narrative content of LIS is an unlikely source of the shift. However, other response categories with significant gender shifts, namely five (negatively valenced personality) and nine (cars), did contain responses that were arguably influenced by the narrative content of LIS.

The percentage of females represented in the negatively valenced personality category went from a pre-game value of 0% to a post-game value of 33.33%. Post-game responses in the negatively valenced personality category containing mentions of illicit marijuana use were observed, an act which, as previously discussed, played a central role in the narrative content in LIS. The gender breakdown of the cars response category also switched from a 100% male population in the pre-game survey to a 100% female population in the post-game. Post-game responses in the cars category containing mentions of US students driving themselves to school were also observed, an act which was alluded to when Max meets Warren who is standing in front of his car in the student parking lot at Blackwell Academy. Although the source of the significant gender shifts in responses is outside the scope of the analyses described above, the specific responses in these significantly shifted categories were arguably affected by the in-game content of LIS.

7.4 Analysis of perception of media sources

For the pre-game survey question asking for which media sources the participants used to inform their impressions on US culture, 10 response categories were identified and are summarized in Figure 7.13. The gender breakdown for each response category is summarized in Figure 7.14. Compared to the gender percentage for the total dataset (males=65.39%, females=34.62%), there were a number of categories with significant gender-group skews. Response categories with higher male to female ratios included: 2) experience (e.g., media consumed while abroad) (males=90.91%), movies (males=76.92%), impressions or stereotypes gathered from general sources (e.g., I heard it from somewhere) (males=80%), and classes and/or textbooks (males=100%). Response categories with higher female to male ratios included: 3) TV (females=46.88%), and news articles (females=55.56%).

Figure 7.13 Perception of valid media sources category breakdown (pre-game)

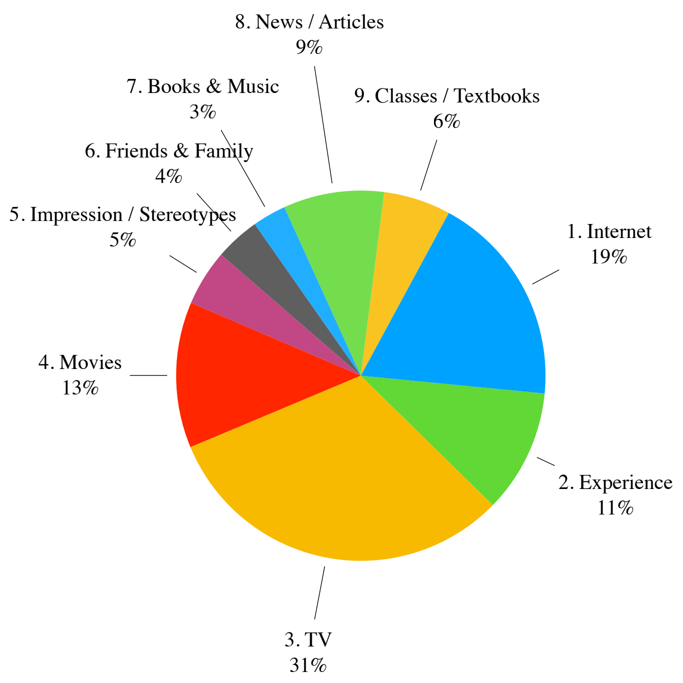
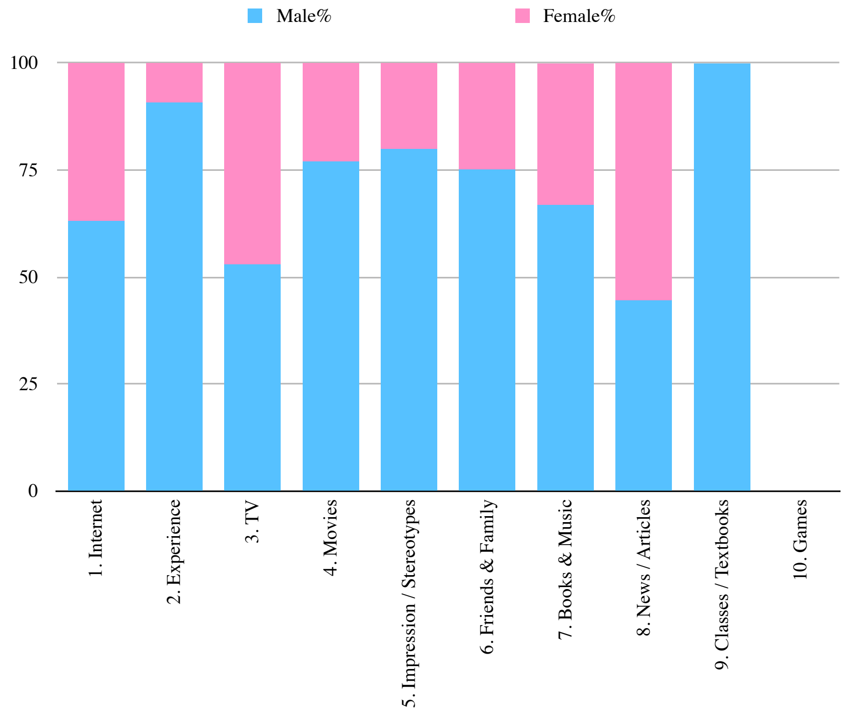


Figure 7.14 Perception of valid media sources gender breakdown (pre-game)



The post-game survey response category breakdown for the perception of valid media sources is summarized in Figure 7.15 and the gender breakdown in Figure 7.16. Paired, two-tailed *t*-tests revealed a highly significant increase (p=0.000053) in the percentage of responses which cited digital games as being valid media sources to inform the participants' impressions on US culture (category number ten) from a pre-game value of 0% to a post-game value of 13.27%. The significant change in the percentage of students who listed digital games as a valid media source to inform their impressions on US culture provides strong evidence to support the argument that the narrative content of digital games such as LIS may be used as pedagogical tools to address the cultural aspects of education in SLA contexts. No other significant changes were observed. There were, however, two significant gender shifts that occurred in response categories six (friends and family) and seven (books and music), whereby the percentage of females in each response category increased from pre-game values of 25% and 33.33% to post-game values of 60% and 66.67% respectively. The gender shifts in response categories six and seven are, however, outside the scope of the present analysis, and it is therefore difficult to argue that the narrative content of LIS had any impact whatsoever on the above-mentioned gender shifts.

Figure 7.15 Perception of valid media sources category breakdown (post-game)

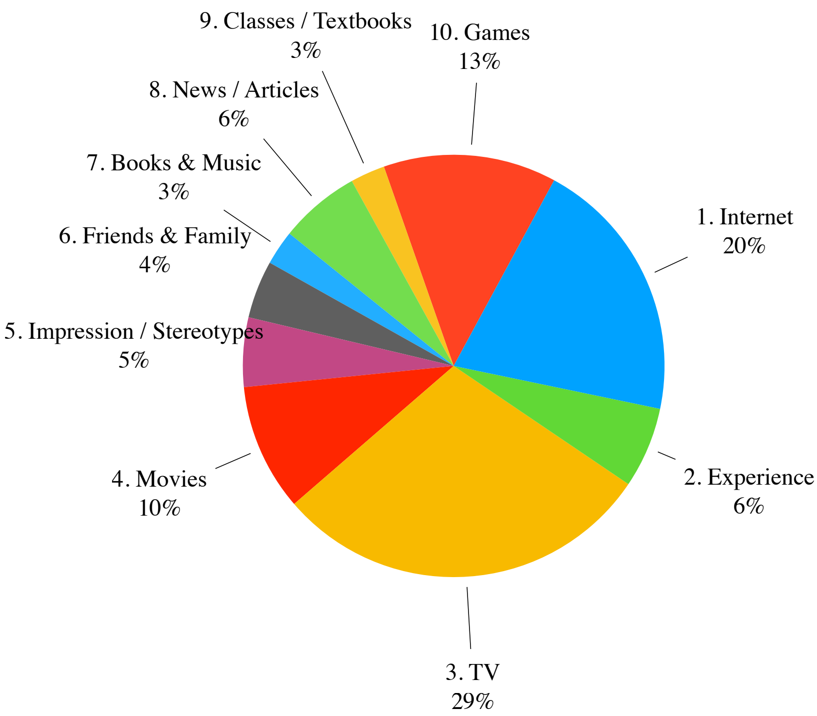
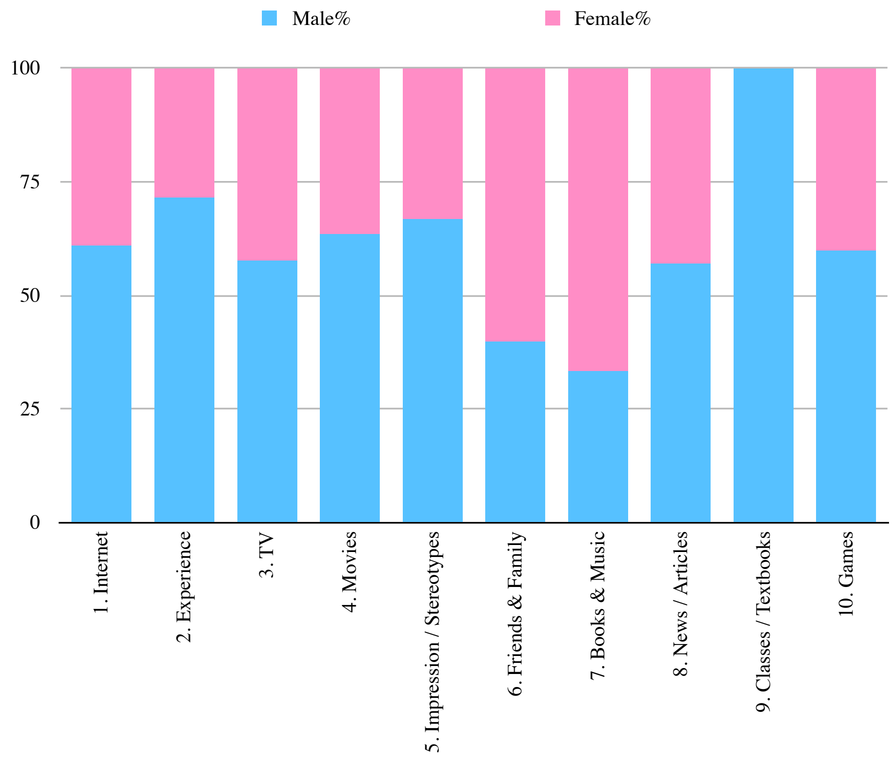


Figure 7.16 Perception of valid media sources gender breakdown (post-game)



7.5 Summary of perceptions of US culture

This chapter presented the results of the qualitative pre-/post-test survey data to analyze how the participants' perceptions of US culture were influenced by the LIS gaming intervention. Significant pre to post-game shifts were observed in all categories: perception of US culture (increase in answers mentioning crime and drugs), perception of the US educational system (decrease in answers mentioning the grade structure), perception of US students (increase in answers mentioning communication skills), and perception of valid media sources (increase in answers mentioning digital games). These shifts in perception are attributable directly to the narrative content of LIS and highlight the need for DGBL curricula to address the accuracy of cultural portrayals in digital games. The implications of these results are analyzed in further detail in the following chapter.

Chapter 8 Conclusions

The following chapter will begin by reviewing and answering each RQ based on the evidence from the data analyses presented in the results chapters. After addressing each RQ with evidenced-based answers, the following sections will frame a discussion of the implications for the use of single-player RPGs such as LIS in foreign language classrooms.

8.1 Evidence-based answers to research questions

8.1.1 Evidenced-based answer to research question one

RQ1 is as follows: 1) How effective is the LIS gaming intervention at eliciting TL vocabulary acquisition and retention when used in conjunction with pre-determined testing protocols relative to a control group? The LIS gaming intervention was shown to be more effective at eliciting vocabulary acquisition in the immediate tests relative to the control intervention (see Table 4.9 and Figure 4.40). Although the vocabulary test scores in the LG group were moderately higher after corrective feedback in the delayed tests (see Figure 4.41), the differences in score relative to the LC group were only moderately significant. However, the LIS gaming intervention was shown to elicit long-term retention of the vocabulary in the OSD vocabulary tests relative to control (see Figure 4.42). Furthermore, rates of vocabulary acquisition and retention were found to be significantly bolstered through multiple exposures to the LIS gaming intervention (Table 4.3).

8.1.2 Evidence-based answer to research question two

RQ2 is as follows: How do the gameplay mechanics in the smartphone application of LIS affect vocabulary acquisition and retention of TL vocabulary? Using the IQA from both LG and LC groups, specific gameplay mechanics from LIS were identified as being potentially conducive for vocabulary acquisition/retention, namely linguistically down-regulated action sequences, mandatory in-game tasks with corrective feedback, and the multimodal presentation of vocabulary (see Table 5.7, Table 5.1, Table 5.2, and Table 5.3). Furthermore, the positive correlation identified between comprehension test scores and vocabulary test scores may implicate comprehension-conducive gameplay mechanics which employ corrective feedback as being particularly conducive to vocabulary acquisition and retention (Figure 5.9).

8.1.3 Evidence-based answer to research question three

RQ3 is as follows: How will the LIS gaming intervention correlate to students' perceptions of and attitudes towards English smartphone gaming? Several factors directly related to smartphone gaming attitudes/perceptions were identified which resulted in significant positive and negative correlations to vocabulary test scores (see Table 6.12). In particular, the following factors were found to have significant correlations to vocabulary test scores in the LIS gaming intervention: gaming location (see Table 6.2), previous English smartphone gaming experience (see Table 6.3), previous LIS gaming experience (see Table 6.4), motivation to play (see Table 6.6), comprehensibility/difficulty of gameplay content (see Table 6.7 and Table 6.10), perception of improvement in English (see Table 6.11), perception of the efficacy of smartphone gaming as a language learning intervention (see Table 6.5 and Table 6.9), and overall level of smartphone engagement (see Table 6.1 and Table 6.8).

8.1.4 Evidence-based answer to research question four

RQ4 is as follows: How will the depictions of US culture in the narrative content of LIS affect how the students inform their impressions on US culture? Shifts in perceptions towards aspects of US culture were directly tied to the narrative content in LIS. In particular, the participants' perceptions of US culture were affected by the LIS narrative content featuring crime and drugs (Figure 7.1 and Figure 7.3), high school environments (Figure 7.5 and

Figure 7.7), and in-class pedagogical style (see Figure 7.9 and Figure 7.11). Furthermore, the participants viewed the gaming intervention as a valid media source to inform these opinions (Figure 7.13 and Figure 7.15).

8.2 Summary and classroom implications of analysis

8.2.1 In-game and out-of-game corrective feedback

Considering the results from the vocabulary testing protocols in both LG and LC groups (see Table 4.9), it can be stated with confidence that the LIS gaming/testing intervention was more effective at eliciting vocabulary acquisition and long-term retention than the control testing intervention. These results suggest that the gameplay mechanics of LIS in conjunction with the classroom interventions (i.e., testing interventions and post-game content discussions/student oral presentations) were more effective in eliciting long-term vocabulary retention than simple testing with corrective feedback alone. Of further note are the results from the immediate testing protocols, whereby the conclusion may be drawn that the gameplay mechanics of LIS were highly effective at eliciting vocabulary acquisition without the aid of any testing/classroom interventions relative to the control intervention. With that being said, however, considering the well-researched supposition presented in chapter three that corrective feedback may elicit long-term retention when presented after a period of inferential processing (Li, 2010), the use of testing/classroom procedures which provide said corrective feedback in conjunction with the gaming procedure may be beneficial in eliciting long-term retention. This supposition is also well supported in the quantitative analysis results, whereby the LC group saw roughly equivalent gains in the delayed post-test protocol (i.e., after corrective feedback) but did not retain these gains as well as the LG group (see Table 4.9). Therefore, if the goal of the gaming intervention in the language classroom is to elicit long-term retention of in-game vocabulary, then testing/classroom protocols which provide corrective feedback are recommended.

Corrective feedback protocols used in conjunction with a gaming intervention in the language classroom may also confer benefits on the acquisition/retention of in-game vocabulary from gameplay sequences which do not incorporate mechanics conducive to vocabulary acquisition/retention. The IQA results presented in chapter five clearly show that certain vocabulary experience higher immediate gains than other vocabulary (see Table 5.1, Table 5.2, and Table 5.3). Indeed, some of the immediate gains in the IQA results even outperform the corrective feedback protocols, but these immediate gains were not observed for all vocabulary test items. In particular, vocabulary from linguistically down-regulated action sequences, mandatory in-game task sequences, and vocabulary with multimodal presentations (e.g., "lighthouse") were all associated with higher immediate gains from gaming-only procedures than other vocabulary test items. However, the use of the corrective feedback protocol still elicited more vocabulary acquisition in the LG group relative to the LC group (see Table 5.7), suggesting that corrective feedback is important to bolster the vocabulary acquisition rates from gaming interventions. In other words, the corrective feedback protocols used in the present study helped the LG group acquire in-game vocabulary that were perhaps more difficult or less influenced by acquisition-conducive gameplay mechanics. These results add further weight to the supposition that long-term retention of vocabulary acquired through in-class gaming interventions may be supported through the use of corrective feedback protocols.

8.2.2 Multiple exposures to the LIS gaming intervention

Perhaps the most notable finding from the quantitative analysis results was the highly efficacious advantage of multiple play-throughs presented in chapter four (see Table 4.3). While it remains outside the scope of the present study to determine whether similar advantages would have been evident with multiple occasions of participation in the LC group, considering the stronger vocabulary acquisition/retention effect observed in the LG group, it may be hypothesized that multiple play-throughs in the LG group would be more efficacious than multiple control tests. Considering the classroom implications of the advantage of multiple play-throughs, designing a DGBL curriculum which makes use of multiple play-throughs of the same gaming material is recommended. Although there are a number of ways to accomplish designing a curriculum to allow for multiple play-throughs of the gaming intervention, considering the benefits to mobility offered by smartphone gaming, utilizing smartphones as the gaming platform would allow instructors to outsource multiple play-through gaming sessions as homework, thus freeing up in-class time for primary gaming sessions and/or other game-based classroom interventions.

8.2.3 Gaming location

The significant difference in gaming location on the vocabulary test averages presented in chapter six (see Table 6.2) is another factor to consider when implementing a DGBL curriculum in a foreign language classroom, especially when utilizing mobile gaming platforms such as smartphones. While the vocabulary test averages were slightly higher from gaming sessions played at the participants' homes during the period of mandatory online learning during the Covid-19 pandemic, so too were the vocabulary test averages from gaming sessions conducted in the classroom relative to gaming sessions conducted while out (e.g., at a café or in transit). Of particular concern for any DGBL curriculum outsourcing smartphone gaming interventions as homework is the potentially detrimental impact that playing smartphone games in distracting environments such as at cafés or riding on the train may have on the overall outcome of the gaming intervention itself. Furthermore, as the data from a number of study participants from the online learning period had to be removed due to non-compliance with the English-only gaming rule stipulated at the start of the study program (see Table 3.1), the lack of supervision over how the students are engaging with the game may have a detrimental impact on the overall outcome of the gaming intervention. It is therefore recommended that at-home smartphone gaming sessions be limited to multiple play-through sessions for review and that the primary gaming sessions be undertaken in the classroom setting.

8.2.4 Prior English gaming experience

In chapter six, another factor which was found to influence the overall efficacy of the LIS gaming intervention was prior experience gaming in English (Table 6.3) and, as would be expected, prior experience playing LIS (especially in English, but also in Japanese or another language as well) (see Table 6.4). Although the claim that gaming in English will increase EFL proficiency is beyond the scope of this study, it was demonstrated in the results of the quantitative analysis that students who had experience playing games in English were more likely to perform better on the vocabulary tests. Other than playing digital games in English, experience with playing LIS in a language other than English was also moderately correlated with higher vocabulary test averages than students with no prior English gaming experience or experience playing LIS (see Figure 6.15 and Figure 6.16). Considering that familiarity with the mechanics and/or narrative content of a digital game may have a beneficial influence on the overall outcome of a DGBL curriculum, it may be beneficial to choose a digital game with which a large percentage of the students have had experience playing before. With that being said, when gaming interventions are used in SLA contexts, digital titles with gameplay mechanics conducive to SLA should be prioritized over general familiarity of the game within the student population. As an example, many students may be familiar with the smartphone title *Candy Crush Saga* (King, 2012), however the lack of SLA-conducive gameplay mechanics in such highly linguistically down-regulated smartphone puzzle games would arguably have negligible impact on linguistic development. Therefore, balancing student familiarity with SLA-conducive gameplay mechanics in any digital game title chosen for a DGBL curriculum may achieve the most beneficial outcome.

8.2.5 Correlation analyses

The correlation analyses presented in chapters five and six revealed a number of factors which were associated with better/worse vocabulary acquisition/retention rates from the LIS gaming intervention (see Table 6.12). As there were no universal measures which could have been employed to judge students' English proficiency (e.g., TOEIC scores), the comprehension tests were designed to not only gauge how much in-game narrative content the participants understood from each of the three gaming sessions but also their general level of English proficiency. The correlation analyses revealed a significant correlation between the vocabulary and comprehension tests (see Figure 5.9, Figure 5.10, and Figure 5.11). These results suggest that generally higher levels of English proficiency were associated with better comprehension and thus better rates of vocabulary acquisition/retention. Furthermore, games with corrective mechanics conducive to eliciting narrative comprehension may yield better results in SLA contexts.

8.2.6 Gender-specific correlation factors

Other factors correlated with better rates of vocabulary acquisition/retention were more gender specific. Males who perceived that they understood the in-game content well were more likely to perform well on comprehension tests relative to females (see Table 6.7), which in turn was correlated to better performance on vocabulary tests (see Figure 5.10 and Table 6.12). Looking at the qualitative data gathered in the post-game survey, one factor that was identified as an obstacle to in-game content comprehension were game-specific mechanics such as controls (see Figure 6.42 and Figure 6.43). Therefore, engagement and familiarity with smartphone gaming in contexts outside the study may have played a role in males perceiving that they understood the in-game content more (see Figure 6.14 and Figure 6.16). The perception of in-game content comprehension was also positively correlated with the perception of English growth (see Table 6.7), which again was associated with smartphone engagement and perception of the efficacy of smartphone gaming as a language learning intervention (see Table 6.12).

Females, just as with males, who performed well on the comprehension tests were more likely to perform well on the vocabulary tests (Figure 5.11). However, relative to the male sample, females who perceived the in-game English to be difficult did worse on the comprehension/vocabulary tests (see Table 6.10). In contrast to these findings, females who performed better on the comprehension/vocabulary tests were more likely to perceive a benefit from the gaming intervention (see Table 6.11), which in turn was associated with better performance on the vocabulary/comprehension tests (see Figure 6.76 and Figure 6.78). These positive correlations between perception of English growth and better performance on the vocabulary/comprehension tests were negative correlations in the male sample (see Figure 6.75 and Figure 6.77), which provides evidence that the females in this study may have valued the performance on the testing protocols over the experience of English gaming relative to males. Furthermore, females who found the in-game English to be difficult were more likely to perceive smartphone gaming as an efficacious language learning tool (Figure 6.65 and Figure 6.66). Just as with males, however, females who perceived a linguistic benefit from the LIS gaming intervention were also more likely to engage in English smartphone gaming in the future (Table 6.8).

Ultimately, the common path to better performance on the vocabulary tests for both genders was through comprehension. However, each gender took different paths to positive comprehension correlations. Males valued comprehending the in-game content more whereas females valued the perception that better performance on the testing protocols was leading to greater gains in English proficiency. Considering the classroom implications of these results, utilizing classroom interventions such as in-class discussions, projects, or assignments to supplement the students' comprehension of the narrative content of the gaming intervention may in turn elicit better long-term retention of any vocabulary acquired during gameplay.

8.2.7 Motivation to play

In chapter six, a confounding factor concerning the positive correlation between vocabulary and comprehension scores was a negative correlation associated with males who were not motivated to continue playing the game (see Table 6.6, Figure 6.35). In general, the enjoyment/motivation factor of playing LIS demonstrated similar positive correlations between vocabulary and comprehension scores (see Table 6.6). However, non-motivated males demonstrated an inverse relationship between vocabulary and comprehension scores (see Figure 6.35), meaning that non-motivated males who scored higher on the comprehension tests subsequently scored lower on the vocabulary test scores. In addition to the argument put forth in the study that English proficiency and general comprehension of in-game narrative content leads to better acquisition rates, these acquisition rates may also be impacted by how much the students enjoy playing the game (especially for males). Although purely speculative, it may be the case that the non-motivated males in this study with higher levels of English proficiency were able to play through the game quickly, thus by-passing the gameplay mechanics potentially conducive to vocabulary acquisition/retention such as corrective feedback mechanics from mandatory in-game tasks, negotiating the meaning of unknown vocabulary with the in-game environment, and using the time reversal mechanics during linguistically down-regulated action sequences exposing the player to repeated dialogue sequences. With these results in mind, considering the ludic value of any digital game title chosen for use in a DGBL curriculum may benefit the overall pedagogical outcome of said gaming intervention.

8.2.8 Shifts in perceptions regarding smartphone gaming

The LIS gaming intervention presented in this study influenced more than vocabulary acquisition and long-term retention. Chapter six presented how the students' perceptions regarding smartphone gaming itself was influenced by the LIS gaming intervention. Shifts in perception from pre to post-study were observed in how willing students were to engage with smartphone gaming in the future as well as in how efficacious a language learning tool the students perceived smartphone gaming to be after exposure to the LIS gaming intervention. Many of these shifts were specific to gender. Although there were no large shifts in correlation from pre- to post-study in how much males rated their perception of smartphone gaming efficacy, a moderate positive shift was observed in males who perceived a benefit to their English proficiency by participating in the LIS gaming intervention (see Table 6.9 and Figure 6.68). The perception of English growth for males was also found to be correlated to how well they comprehended the in-game narrative content (see Table 6.7 and Figure 6.39). In other words, males who comprehended the in-game content were more likely to perceive that their English skill benefited from playing LIS, which, in turn, was associated with an increase to the perceived efficacy of smartphone games as a language learning tool. These results also have implications for classroom interventions, whereby making use of in-class protocols to bolster comprehension of the in-game narrative content may increase the perceived benefit to the students' English growth and encourage the perception of smartphone games as a beneficial tool for language learning.

For females, there was a notable shift in how efficacious they perceived smartphone gaming to be as a language learning intervention. More specifically, a moderate number of females perceived smartphone gaming to be an efficacious language learning tool post-study relative to pre-study values (see Table 6.10). A correlation to this shift in smartphone gaming efficacy perception was observed in how difficult females found the game to be, whereby the more difficult the game was perceived to be, the more efficacious smartphone gaming was perceived to be (see Figure 6.65 and Figure 6.66). Therefore, when choosing a digital game title for use in a DGBL curriculum in a foreign language classroom, the more difficult the narrative content of the game, the more efficacious female students may perceive the gaming intervention to be. However, in order to bolster how much benefit a gaming intervention is perceived to have on English language skills, in-class protocols discussing in-game narrative content would again be beneficial to the overall outcome of the gaming intervention.

Considering the benefits to vocabulary acquisition/retention associated with prior English smartphone gaming experience (see Table 6.3), encouraging students to continue to play smartphone games in English may also confer more benefits in the development of their English language proficiency. Although there were no notable shifts in both genders' intentions to engage with smartphone games from pre- to post-study, there was a shift in a correlational relationship in the female sample. More specifically, females who perceived their English skill to have benefited from the LIS gaming intervention were more likely to engage with smartphone games post-study (see Table 6.9, Figure 6.45). Furthermore, females who performed better on the vocabulary/comprehension tests were more likely to engage with smartphone gaming post-study (see Figure 6.50, Figure 6.51, Figure 6.54, and Figure 6.55). Therefore, in order to encourage students to continue engaging with smartphone games after completion of an in-class gaming intervention, choosing a digital game title with relatively difficult narrative content, and supplementing with in-class protocols to bolster comprehension may in turn yield better long-term pedagogical results through continued independent engagement with smartphone gaming used as a language learning tool.

8.2.9 Shifts in perceptions regarding US culture

As SLA involves more than the acquisition of TL vocabulary (Haastrup & Henriksen, 2001; Scovel, 1994), the potential impact that a gaming intervention using a single-player RPG like

LIS may have on the perception of the TL culture (in the case of LIS, US culture) was also analyzed. In chapter seven, statistically significant shifts in the perception of US culture were identified, all of which were arguably the direct result of the portrayal of US culture featured in the narrative content of LIS. Three statistically significant shifts were identified regarding US culture: 1) crime and drugs (see Figure 7.1 and Figure 7.3), 2) educational style/structure (see Figure 7.5 and Figure 7.7), and 3) communication style (see Figure 7.9 and Figure 7.11). With regard to crime and drugs, the majority of responses contained mentions of gun violence and marijuana use in the US, topics which LIS directly addresses with the school shooting scene in the girls' bathroom at Blackwell Academy and with Chloe's open use of marijuana in her bedroom. Furthermore, the in-game dialogue between Nathan and Chloe during the school-shooting scene in the girls' bathroom at Blackwell Academy revolves around the topic of drugs and money. This topic is further elaborated upon during conversations between Chloe and Max later on in the game. The shifts in the perception of educational style/structure were arguably directly influenced by the in-game depictions of Blackwell Academy and Mr. Jefferson's photography lecture. More specifically, a large portion of responses mentioned how US pedagogical approaches make use of communicative discussions and active participations in lectures. Finally, a number of students mentioned that US students are highly communicative, both in lessons and with each other. As the narrative content and gameplay mechanics in LIS revolve around interpersonal dialogue and script/narrative choice mechanics, the nature of these responses are not surprising.

Perhaps even more notable than the statistically significant shifts in perceptions of US culture was the inclusion of digital games as a valid media source to inform these impressions from the post-game survey (see Figure 7.13 and Figure 7.15). Considering that the narrative content of smartphone games has the ability to influence impressions of TL culture, it is vital that any in-class protocols used in conjunction with gaming interventions address the accuracy of any cultural portrayals, whether directly related to the TL culture or not. For better or for worse, the statistically significant shifts in US culture influenced by the narrative content of LIS are largely accurate, albeit exaggerated for narrative effect. Unfortunately, the US does have a problem with gun violence (Warnick, Kim, & Robinson, 2015) and drug use (Kaliszewski, 2019) in schools, especially when compared to Japan (Kopel, 1993; Shimane & Wada, 2007). Likewise, the perception that pedagogical methods in US schools and US students are more communicative relative to Japan may have a basis in reality (Hosoki, 2011; Schneider, 2013). For every generalization about a specific culture that can be made, however, counterarguments may also be made. Therefore, in-class discussions about the accuracy of portrayals made about any culture within the narrative content of a digital game title are recommended.

8.3 Limitations of the LIS gaming intervention in formal academic settings

Considering the difficulties involved with integrating time-intensive digital games such as single-player RPGs into a time-restrictive environment such as a foreign language classroom, there were a number of limitations in the gaming intervention presented in this study. First and foremost, as was just mentioned, the length of the full version of LIS was far too long to use in the time allotted to the courses in which this study program was conducted. As outlined in the methodology section, each course in which this study program was conducted met once weekly for 90 minutes over a semester of 15 weeks. Considering that the full game of LIS can take up to 16 hours to complete, without outsourcing significant portions of gameplay as homework, it would have been difficult to incorporate the entire game into one semester's worth of courses while attending to other curriculum obligations (i.e., coursework other than smartphone games). Furthermore, should the full version have been used in the study, the cost of the game for the LG group would have presented a significant financial barrier to the author or would have had to have been borne by the participants themselves. With these time and financial constraints in mind, the study only utilized the first chapter of LIS, which was available for free as a preview of the full version of the game. However, limiting the study to the first chapter, played over a course of three weeks, may have limited the efficacy of the game as a language learning tool.

Although one of the benefits to using a single-player RPG as a medium of instruction in formal academic environments over MMORPGs is the more linear nature of the game's narrative content, this is not to say that all players follow exactly the same trajectory when playing through a single-player RPG such as LIS. The trajectory that an individual player follows through online games such as MMORPGs can be highly unique, which make them more difficult to utilize in conjunction with in-class protocols (e.g., vocabulary tests, comprehension tests, or narrative discussions) because most of the students enrolled in the course will not have experienced the same dialogue, in-game events, and narrative content. In contrast to the unique gameplay trajectories players follow in online genres of digital games such as MMORPGs, single-player RPGs will usually incorporate mandatory sequences of gameplay that occur in a pre-determined sequence. While the more linear gameplay trajectories players follow in single-player RPGs make the genre more conducive to integration in a DGBL curriculum, players of single-player RPGs still follow unique gameplay trajectories in between mandatory narrative sequences as they control which NPCs they communicate with, when to initiate narrative sequences, and what optional gameplay mechanics (e.g., side quests) to participate in. Considering that one of the goals of this study was to analyze how interaction with the gameplay mechanics of single-player RPGs affects the acquisition of in-game vocabulary, then certainly allowing students to play the game as they wanted, when they wanted, may have produced different pedagogical outcomes.

However, due to the time constraints of the study program, all students had to play the game on exactly the same schedule.

A particular knock-on effect of constraining gameplay to a universal in-class schedule is that the spacing effect for vocabulary acquisition and retention will be affected. For example, players who take their time playing through the game, talk to all the NPCs, participate in all the optional side quests, and, in the case of LIS, re-play certain narrative sequences to uncover the effects of certain script/narrative choices on the outcome of the game, then the spacing effect may induce better long-term retention for any vocabulary acquired through gameplay. The opposite also rings true, whereby players who rush to complete the game, only talk to the required NPCs, do none of the optional side quests, and do not take advantage of the time reversal mechanics to redo specific script/narrative choices, then the spacing effect would arguably not be as effective at inducing long-term retention of vocabulary acquired through gameplay for those players. Therefore, a more precise measurement of how gameplay style affects vocabulary acquisition/retention would shed more light on the effectiveness of single-player RPGs as language learning tools. With that being said, the time constraints of the present study made it difficult to incorporate for different playing styles as all students had to complete certain gameplay sequences within a pre-determined amount of time in order to participate in the study program.

With the possibility for different gameplay styles in mind, any vocabulary encountered outside the mandatory gameplay sequences in the LIS gaming intervention (e.g., through optional side quests) were impossible to measure from pre- to post-game. In other words, in order to measure the effect that playing LIS had, it was necessary to test the students' knowledge of a pre-determined list of vocabulary from mandatory gameplay sequences before initiating gameplay. However, the testing method employed to measure the efficacy of LIS as a tool to improve vocabulary acquisition in formal EFL contexts is necessarily limited to vocabulary encountered in mandatory gameplay sequences. Unfortunately, this testing method does not provide a full understanding of exactly which vocabulary students are acquiring from exactly which gameplay sequences. Case in point, it could be argued that the vocabulary encountered through optional gameplay sequences which the player initiates him/herself are more conducive to acquisition and long-term retention. The testing method used in this study to measure the pre- to post-game growth of vocabulary therefore does not account for the unique gameplay trajectories that players follow.

Finally, the three-week gameplay schedule employed in this study may not account for the time necessary for long-term consolidation of the vocabulary acquired through gameplay. As research in the field of vocabulary acquisition from digital games has suggested, long-term consolidation of vocabulary acquired through gameplay may not occur until two to six months of continual engagement with the digital game (Cobb & Horst, 2011; Wrobetz, in press). The three-week gaming schedule used in the present study is, by comparison, significantly shorter and may not capitalize on the long-term benefits of vocabulary consolidation through sustained gameplay. However, as was described above, the semester time constraints in conjunction with the financial barriers involved in furnishing all study participants with full versions of LIS, the three-week gameplay period was determined to be the ideal amount of time to play through the first free-to-play chapter of LIS.

8.4 Future considerations

Although the efficacy of the LIS gaming intervention presented in this study in SLA contexts has been established, there are still a number of lines of inquiry that can be made to potentially strengthen the findings presented in this study and to further elaborate upon the mechanisms of TL vocabulary acquisition at play in LIS. First and foremost, the possibility of the pre-game testing protocol to be eliciting a priming effect on subsequent testing protocols should be explored. In particular, whether or not the same levels of acquisition/retention would occur without a pre-game testing protocol is a potential avenue of research. Furthermore, the question of whether long-term retention was significantly affected by the corrective feedback protocols, or whether similar rates of long-term retention would have been achieved without the corrective feedback protocols may also be investigated. Similarly, how the in-class protocols (e.g., in-class discussions of narrative content), testing schedules, and corrective feedback protocols affected acquisition/retention rates remains, as of yet, unknown. Finally, the question of whether the results observed in the LIS gaming intervention presented in this study are unique to this particular digital game title, or whether similar gaming/testing protocols with other digital game titles/media (e.g., televised dramas with subtitles) would elicit similar results should also be explored.

8.5 Conclusion

The LIS gaming intervention presented in this study was shown to be highly efficacious at eliciting vocabulary acquisition and long-term retention and provides evidence for the efficacy of single-player RPGs as foreign language learning interventions in formal academic environments. With the exception of a few software bugs associated with specific iPhone models, the smartphone gaming platform chosen for the LIS gaming intervention made the integration of this DGBL curriculum virtually seamless as all necessary hardware was already in possession by the study participants. Although the possibility existed for there to be a student without access to a LIS application-compatible smartphone or computer, the issue never arose across 172 students and over the course of three years. Furthermore, the choice of the smartphone gaming platform proved to be a highly resilient pedagogical method of instruction when courses switched to online formats during the Covid-19 pandemic. Indeed, the sudden and unexpected shift to online learning during the Covid-19 pandemic opened up potentially vital lines of research by unearthing a gaming location effect, whereby gaming at home was shown to be the most conducive environment for vocabulary acquisition and retention.

Outside of the efficacy for vocabulary acquisition and retention, the unique narrative content of LIS may have exposed students, for better or worse, to aspects of US culture that they might not have been familiar with before, some of which would be difficult to broach in EFL classrooms such as gun violence and drug use in US culture. The moderately accurate portrayal of a US school setting may have also provided a context of language use that is often missing from textbook-driven curricula, namely the authentic depiction of nonstandard English and general communication style within the context of a narrative structure. The gameplay mechanics which required students to comprehend the context of the language being used in order to solve in-game tasks incorporating the quasi-productive script-choice mechanics arguably made the game more difficult than a passive lecture, whereby students do not need to demonstrate comprehension in order to complete the lesson. However, this particular aspect of difficulty, of needing to interact with the context of the language being used in very specific ways to demonstrate a certain level of comprehension, actually compelled a number of students (particularly females) to change their minds about the efficacy of smartphone games as a language learning tool and to seek out more smartphone games in the future.

While MMORPGs have rightfully received a lot of attention as potential sources of SLA instruction, the results of this study show that single-player RPGs such as LIS with gameplay mechanics conducive to SLA hold a lot of promise as language learning interventions. Indeed, when considering the constraints of EFL curricula in formal academic environments, single-player RPGs may be even more conducive than MMORPGs due in part to their much more linear gameplay trajectories. The similarity in gameplay trajectories that players follow in single-player RPGs arguably make the in-game narrative content much more conducive to in-class instruction, whereby comprehension-building activities (e.g., in-class discussions, oral presentations, or coursework) may be designed around the higher frequency of mandatory and sequential gameplay progression relative to other digital game genres such as MMORPGs. Furthermore, the largely pre-determined story and gameplay trajectories may make outsourcing gameplay as homework more feasible, whereby more in-class time may be devoted to comprehension-building exercises.

Finally, with specific regard to the Covid-19 pandemic which prevented students studying foreign language from hopping on a plane and experiencing first hand another language, another culture, another way of life, and ultimately another way of making sense of our shared reality, digital environments such as single-player RPGs like LIS may serve as a stop-gap measure to keep students interested and motivated to keep learning foreign languages during a time which kept most of the world isolated from one another. At the very least, from a purely academic perspective, the LIS gaming intervention presented in this study was shown to be more effective at eliciting long-term retention of TL vocabulary relative to simple testing and corrective feedback. If one of the goals as foreign language instructors is to ensure that our students retain the vocabulary they learn in class long after the class is over, then single-player RPGs should not be overlooked as potentially powerful mediums of foreign language instruction.

References

Aghlara, L., & Tamjid, N. H. (2011). The effect of digital games on Iranian children's vocabulary retention in foreign language acquisition. *Procedia-Social and Behavioral Sciences*, 29, 552-560. Elsevier. https://doi.org/10.1016/j.sbspro.2011.11.275

All, A., Castellar, E. P. N., & Van Looy, J. (2016). Assessing the effectiveness of digital game-based learning: Best practices. *Computers & Education, 92*, 90-103. Elsevier. https://doi.org/10.1016/j.compedu.2015.10.007

Algeo, J. (1999). *Cambridge History of English Language*. Cambridge University Press.

Alsop, T. (2021, February 18). *How many people have access to a computer 2019*. Statista. https://www.statista.com/statistics/748551/worldwide-households-with-computer/

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *The International Review of Research in Open and Distributed Learning*, *4*(2). AU Press. https://doi.org/10.19173/irrodl.v4i2.149

An explanation of L33t speak. (2002). H2g2. https://h2g2.com/edited_entry/A787917

Aponte, M. V., Levieux, G., & Natkin, S. (2009). Scaling the level of difficulty in single player video games. In S. Natkin & J. Dupire (Eds.), *International Conference on Entertainment Computing* (pp. 24-35). Springer. https://doi.org/10.1007/978-3-642-04052-8_3

Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research*, *32*(2), 266-276. Oxford University Press. https://doi.org/10.1086/432236

Auble, P. M., & Franks, J. J. (1978). The effects of effort toward comprehension on recall. *Memory, & Cognition*, *6*(1), 20-25. Springer. https://doi.org/10.3758/BF03197424

Auble, P. M., Franks, J. J., & Soraci, S. A. (1979). Effort toward comprehension: Elaboration or “aha”?. *Memory, & Cognition*, *7*(6), 426-434. Springer. https://doi.org/10.3758/BF03198259

Autio, P., & Takamaa, J. (2020). *How do MMORPGs facilitate language learning for Finnish EFL students?* [Master’s Thesis, University of Jyväskylä]. https://jyx.jyu.fi/bitstream/handle/123456789/69732/URN%3ANBN%3Afi%3Ajyu-202006053989.pdf?sequence=1

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, *4*(5), 316-321. Sage Publishing. https://doi.org/10.1111/j.1467-9280.1993.tb00571.x

Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning*, *57*(1), 35-56. Wiley. https://doi.org/10.1111/j.1467-9922.2007.00398.x

Bates, J. C., & Goodman, E. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*(5-6), 507-584. Routledge. https://doi.org/10.1080/016909697386628

Bernaus, M., & Gardner, R. C. (2008). Teacher motivation strategies, student perceptions, student motivation, and English achievement. *The Modern Language Journal*, *92*(3), 387-401. Wiley. https://doi.org/10.1111/j.1540-4781.2008.00753.x

Birk, M. V., Atkins, C., Bowey, J. T., & Mandryk, R. L. (2016, May). Fostering intrinsic motivation through avatar identification in digital games. In J. Kaye & A. Druin (Eds.), *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2982-2995). Association for Computing Machinery. https://doi.org/10.1145/2858036.2858062

Bisson, M. J., Van Heuven, W. J., Conklin, K., & Tunney, R. J. (2014). The role of repeated exposure to multimodal input in incidental acquisition of foreign language vocabulary. *Language learning*, *64*(4), 855-877. Wiley. https://doi.org/10.1111/lang.12085

Bjork, R. A. (2014). Forgetting as a friend of learning. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger, III (Eds.), *Remembering* (pp. 39-52). Psychology Press. https://doi.org/10.4324/9781315752808

Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning*, *30*(7), 709-725. Routledge. https://doi.org/10.1080/09588221.2017.1356335

Borrás, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal*, *78*(1), 61-75. Wiley. https://doi.org/10.2307/329253

Bostan, B. (2009). Player motivations: A psychological perspective. *Computers in Entertainment (CIE)*, *7*(2), 1-26. Association for Computing Machinery. https://doi.org/10.1145/1541895.1541902

Bowers, P., & London, P. (1965). Developmental correlates of role-playing ability. *Child Development*, *36*(2), 499-508. Wiley. https://doi.org/10.2307/1126472

Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics*, 6(1), 60-70. Oxford University Press. https://doi.org/10.1093/applin/6.1.60

Breiner-Sanders, K. E., Lowe, P., Miles, J., & Swender, E. (2000). ACTFL proficiency guidelines-Speaking (Revised 1999). *Foreign Language Annals*, *33*(1), 13-18. Wiley. https://doi.org/10.1111/j.1944-9720.2000.tb00885.x

Breuer, J., & Bente, G. (2010). Why so serious? On the relation of serious games and learning. *Journal for Computer Game Cultur*e, *4*(1), 7-24. https://hal.archives-ouvertes.fr/hal-00692052

Broussard, M. J. S. (2012). Digital games in academic libraries: A review of games and suggested best practices. *Reference Services Review, 40*(1), 75-89. Emerald Publishing. https://doi.org/10.1108/00907321211203649

Bryant, T. (2006). *Using World of Warcraft and other MMORPGs to foster a targeted, social, and cooperative approach toward language learning*. Academic Commons. https://web.archive.org/web/20061013063948/http://www.academiccommons.org/commons/essay/bryant-MMORPGs-for-SLA

Burghardt, G. M. (2011). Defining and recognizing play. In A. D. Pellegrini (Ed.), *The Oxford handbook of the development of play* (pp. 9-18). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195393002.013.0002

Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2008). Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 918-928. American Psychological Association. https://doi.org/10.1037/0278-7393.34.4.918

Bytheway, J. (2014). In-game culture affects learners' use of vocabulary learning strategies in massively multiplayer online role-playing games. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 4*(4), 1-13. IGI Global. https://doi.org/10.4018/ijcallt.2014100101

Bytheway, J. (2015). A taxonomy of vocabulary learning strategies used in massively multiplayer online role-playing games. *CALICO Journal, 32*(3), 508-527. Equinox Publishing. https://doi.org/10.1558/cj.v32i3.26787

Callan, D. E., & Schweighofer, N. (2010). Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory. *Human Brain Mapping*, *31*(4), 645-659. Wiley. https://doi.org/10.1002/hbm.20894

Chapelle, C. A. (2006). Interactionist SLA theory in CALL research. In J.L. Egbert & G.M. Petrie (Eds.), *CALL research perspectives* (pp. 65-76). Routledge. https://doi.org/10.4324/9781410613578

Chavez, M. M. T. (1998). Learner's perspectives on authenticity. *International Review of Applied Linguistics, 36*(4), 277-306. De Gruyter.

Chen, D. (2015). *Gamer perception of language learning and L2 interaction in MMORPGs*. [Master's Thesis, Utrecht University].

Chen, Z. H., Chen, H. H. J., & Dai, W. J. (2018). Using narrative-based contextual games to enhance language learning: A case study. *Journal of Educational Technology, and Society*, *21*(3), 186-198. https://www.learntechlib.org/p/190827/

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., Clifford, R.M., Early, D.M., & Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in prekindergarten. *Child Development*, *81*(5), 1534-1549. Society for Reseach in Child Development. https://doi.org/10.1111/j.1467-8624.2010.01490.x

Choices and consequences (season 1). (2021). Life Is Strange Wiki. https://life-is-strange.fandom.com/wiki/Choices_and_Consequences_(Season_1)

Chotipaktanasook, N., & Reinders, H. (2018). A massively multiplayer online role-playing game and its effects on interaction in the second language: Play, interact, and learn. In B. Zou & M. Thomas (Eds.), *Handbook of research on integrating technology into contemporary language learning and teaching* (pp. 367-389). IGI Global. https://doi.org/10.4018/978-1-5225-5140-9.ch018

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79-122. Sage Publishing. https://doi.org/10.3102/0034654315582065

Clement, J. (2021). *Topic: Mobile gaming market in the United States*. Statista. https://www.statista.com/topics/1906/mobile-gaming/

Cobb, T., & Horst, M. (2011). Does word coach coach words?. *CALICO Journal*, *28*(3), 639-661. Equinox Publishing. https://doi.org/10.11139/cj.28.3.639-661

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. Elsevier. https://doi.org/10.1016/j.compedu.2012.03.004

Coppens, X., Rico, M., & Agudo, J. E. (2013). Using blogs: authentic material and ranking quality for SLA. *The EUROCALL Review*, *21*(1), 20-34. https://doi.org/10.4995/eurocall.2013.10160

Cornillie, F., Jacques, I., De Wannemacker, S., Paulussen, H., & Desmet, P. (2010). Vocabulary treatment in adventure and role-playing games: A playground for adaptation and adaptivity. In S. De Wannemacker, G. Clarebout, P. De Causmaecker (Eds.), *International conference on interdisciplinary research on technology* (pp. 131-146). Springer. https://doi.org/10.1007/978-3-642-20074-8_11

Cornillie, F., Thorne, S. L., & Desmet, P. (2012). Digital games for language learning: from hype to insight?. *ReCALL, 24*(3), 243-256. Cambridge University Press. https://doi.org/10.1017/S0958344012000134

Cornillie, F., Clarebout, G., & Desmet, P. (2012). The role of feedback in foreign language learning through digital role playing games. *Procedia-Social and Behavioral Sciences*, *34*, 49-53. Elsevier. https://doi.org/10.1016/j.sbspro.2012.02.011

Coxhead, A., & Bytheway, J. (2015). Learning vocabulary using two massive online resources: You will not blink. In D. Nunan & J.C. Richards (Eds.), *Language learning beyond the classroom* (pp. 81-90). Routledge.

Creswell, J. W. (1999). Mixed-method research: Introduction and application. In G. Cizek (Ed.), *Handbook of educational policy* (pp. 455-472). Academic Press.

Day, R. R., Bamford, J., Renandya, W. A., Jacobs, G. M., & Yu, V. W. S. (1998). Extensive reading in the second language classroom. *RELC Journal*, *29*(2), 187-191. Sage Publishing. https://doi.org/10.1177/003368829802900211

De Guerrero, M. C., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *The Modern Language Journal, 84*(1), 51-68. Wiley. https://doi.org/10.1111/0026-7902.00052

DeHaan, J., Reed, W. M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, *14*(2), 74-94. University of Hawaii National Foreign Language Resource Center. https://doi.org/10125/44215

Demographics of mobile device ownership and adoption in the United States. (2021, November 23). Pew Research Center: Internet, Science, & Tech. https://www.pewresearch.org/internet/fact-sheet/mobile/

Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, *43*(8), 627-634. American Psychological Association. https://doi.org/10.1037/0003-066X.43.8.627

Desktop Operating System Market Share Worldwide. (2021). StatCounter Global Stats. https://gs.statcounter.com/os-market-share/desktop/worldwide/#monthly-202011-202110

Dickey, M. D. (2007). Game design and learning: A conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development*, *55*(3), 253-273. Springer. https://doi.org/10.1007/s11423-006-9004-7

Dinçay, T. (2012). Why not teach slang in the classroom?. *Dil Dergisi, 155*, 24-34. https://doi.org/10.1501/Dilder_0000000164

Dixon, D. H. (2014). *Leveling up language proficiency through massive multiplayer online role playing games: Opportunities for English learners to receive input, modify output, negotiate meaning, and employ language-learning strategie*s. [Master's Thesis, The University of Utah].

Dixon, D. H., & Christison, M. (2021). L2 gamers' use of learning and communication strategies in massively multiplayer online games (MMOs): An analysis of L2 interaction in virtual online environments. In K.B. Kelch, P. Byun, S. Safavi, & S. Cervantes (Eds.), *CALL theory applications for online TESOL education* (pp. 296-321). IGI Global. https://doi.org/10.4018/978-1-7998-6609-1.ch013

Doolittle, P. E. (1997). Vygotsky's zone of proximal development as a theoretical foundation for cooperative learning. *Journal on Excellence in College Teaching*, *8*(1), 83-103. https://eric.ed.gov/?id=EJ577702

Dörnyei, Z. (1990). Conceptualizing motivation in foreign-language learning. *Language Learning, 40*(1), 45-78. Wiley. https://doi.org/10.1111/j.1467-1770.1990.tb00954.x

Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. Teachers College Press.

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

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, *28*(2), 339-368. Cambridge University Press. https://doi.org/10.1017/S0272263106060141

Ensslin, A. (2012). *The language of gaming*. Basingstoke; Palgrave Macmillan.

Erdener, D. (2016). Basic to applied research: the benefits of audio-visual speech perception research in teaching foreign languages. *The Language Learning Journal*, *44*(1), 124-132. Routledge. https://doi.org/10.1080/09571736.2012.724080

Félix-Brasdefer, J. C. (2018). Role plays. In A. Jucker, K. Schneider, & W. Wolfram (Eds.), *Methods in pragmatics* (pp. 305-332). De Gruyter. https://doi.org/10.1515/9783110424928

Fernandes, P., Leite, C., Mouraz, A., & Figueiredo, C. (2013). Curricular contextualization: Tracking the meanings of a concept. *The Asia-Pacific Education Researcher*, *22*(4), 417-425. Springer. https://doi.org/10.1007/s40299-012-0041-1

Fernández-García, M., & Martínez-Arbelaiz, A. (2002). Negotiation of meaning in nonnative speaker-nonnative speaker synchronous discussions. *CALICO Journal, 19*(2), 279-294. Equinox Publishing. https://doi.org/10.1558/cj.v19i2.279-284

Final Fantasy IV [Video Game]. (1991). Square.

Final Fantasy VII [Video Game]. (1997). Square.

Final Fantasy IX [Video Game]. (2000). Square.

Final Fantasy X [Video Game]. (2001). Square Product Development Division 1.

Flamand, E. (2008). *Deciphering L33t5p34k internet slang on message boards*. [Master's Thesis, Ghent University]. https://lib.ugent.be/fulltxt/RUG01/001/414/289/RUG01-001414289_2010_0001_AC.pdf

Fletcher, D. (2009, July 29). *Hipsters*. Time. https://web.archive.org/web/20090730232530/http://www.time.com/time/arts/article/0,8599,1913220,00.html

Fowler, C. A., Napps, S. E., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory, & Cognition*, *13*(3), 241-255. Springer. https://doi.org/10.3758/BF03197687

Frumuselu, A. D., De Maeyer, S., Donche, V., & Plana, M. D. M. G. C. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education, 32*(B), 107-117. Elsevier. https://doi.org/10.1016/J.LINGED.2015.10.001

Fuente, M. J. (2003). Is SLA interactionist theory relevant to CALL? A study on the effects of computer-mediated interaction in L2 vocabulary acquisition. *Computer Assisted Language Learning, 16*(1), 47-81. Routledge. https://doi.org/10.1076/call.16.1.47.15526

García-Carbonell, A., Rising, B., Montero, B., & Watts, F. (2001). Simulation/gaming and the acquisition of communicative competence in another language. *Simulation & Gaming*, *32*(4), 481-491. Sage Publishing. https://doi.org/10.1177/104687810103200405

Gee, J. P. (2003). *Learning about learning from a video game: Rise of Nations*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.543.3194&rep=rep1&type=pdf

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)*, *1*(1), 20. Association for Computing Machinery. https://doi.org/10.1145/950566.950595

Gee, J. P. (2004). Learning by design: Games as learning machines. *Interactive Educational Multimedia: IEM, 8,* 15-23. Digital Education Observatory; Virtual Teaching and Learning Group. http://www.ub.es/multimedia/iem

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-learning and Digital Media*, *2*(1), 5-16. Sage Publishing. https://doi.org/10.2304/elea.2005.2.1.5

Gilakjani, A. P. (2016). English pronunciation instruction: A literature review. *International Journal of Research in English Education*, *1*(1), 1-6. http://ijreeonline.com/article-1-21-en.html

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1-67. American Psychological Association. https://doi.org/10.1037/0033-295X.91.1.1

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, *40*(2), 97-118. Cambridge University Press. https://doi.org/10.1017/S0261444807004144

Gilmore, A. (2011). "I prefer not text": Developing Japanese learners' communicative competence with authentic materials. *Language Learning*, *61*(3), 786-819. Wiley. https://doi.org/10.1111/j.1467-9922.2011.00634.x

Godwin-Jones, R. (2018). Contextualized vocabulary learning. *Language Learning & Technology*, *22*(3), 1-19. University of Hawaii National Foreign Language Resource Center. https://doi.org/10125/44651

Goh, S. L. (2016). Learning strategies and learner attitudes in the massively multiplayer online role-playing game cube world. *JALT CALL Journal*, *12*(3), 203-216. https://files.eric.ed.gov/fulltext/EJ1125244.pdf

Grand Theft Auto: San Andreas [Video Game]. (2004). Rockstar North.

Grode, J., & Stacy, A. (2015). Authentic materials and project-based learning. In D. Nunan & J.C. Richards (Eds.), *Language learning beyond the classroom* (pp. 171-179). Routledge. https://doi.org/10.4324/9781315883472

Gurwin, G. (2021, March 16). *The history of the Xbox*. Digital Trends. https://www.digitaltrends.com/gaming/the-history-of-the-xbox

Haastrup, K., & Henriksen, B. (2001). The interrelationship between vocabulary acquisition theory and general SLA research. *EUROSLA Yearbook*, *1*(1), 69-78. John Benjamins Publishing Company. https://doi.org/10.1075/eurosla.1.08haa

Haenfler, R. (2013). *Subcultures: The basics*. Routledge.

Habgood, M. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *The Journal of the Learning Sciences*, *20*(2), 169-206. Routledge. https://doi.org/10.1080/10508406.2010.508029

Heidt, J. S. (2020). *"GG" good gaming wordlists: Insight into single-player commercial off-the-shelf games and action-roleplay game vocabulary*. [Doctoral dissertation, Carleton University].

Hill, E. H., & Giammatteo, M. C. (1963). Socio-economic status and its relationship to school achievement in the elementary school. *Elementary English*, *40*(3), 265-270. National Council of Teachers of English. https://www.jstor.org/stable/41385452

Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition*, *38*(1), 102-115. Springer. https://doi.org/10.3758/MC.38.1.102

Hitosugi, C. I., Schmidt, M., & Hayashi, K. (2014). Digital game-based learning (DGBL) in the L2 classroom: The impact of the UN's off-the-shelf videogame, Food Force, on learner affect and vocabulary retention. *CALICO Journal*, *31*(1), 19-39. Equinox Publishing. https://doi.org/10.11139/CJ.31.1.19-39

Honeywood, R. (2007, March). *The Square-Enix approach to localization*. [PowerPoint slides]. http://ubm-twvideo01.s3.amazonaws.com/o1/vault/gdc07/slides/S4672i1.pdf

Horowitz, K. S. (2019). Video games and English as a second language: The effect of massive multiplayer online video games on the willingness to communicate and communicative anxiety of college students in Puerto Rico. *American Journal of Play*, *11*(3), 379-410. The Strong. https://eric.ed.gov/?id=EJ1220304

Hosoki, Y. (2011). English language education in Japan: Transitions and challenges (I). *Kokusai Kankei Bulletin, 6*(1), 199-215. http://id.nii.ac.jp/1265/00000272/

How long is final fantasy IX?. (2021). HowLongToBeat. https://howlongtobeat.com/game?id=3505

How long is final fantasy X?. (2021). HowLongToBeat. https://howlongtobeat.com/game?id=3525.

How long is life is strange?. (2021). HowLongToBeat. https://howlongtobeat.com/game?id=21702

Huang, B. G., & Yang, J. C. (2014, August). The effects of prior knowledge for incidental vocabulary acquisition on multiplayer online role-playing game. In E. Popescu, R.W.H. Lau, K. Pata, H. Leung, & M. Laanpere (Eds.), *International conference on web-based learning* (pp. 98-105). Springer.

Hubbard, P. (1991). Evaluating computer games for language learning. *Simulation & Gaming*, *22*(2), 220-223. Sage Publishing. https://doi.org/10.1177/1046878191222006

Hüseyin, Ö. Z., & Efecioğlu, E. (2015). Graphic novels: An alternative approach to teach English as a foreign language. *Journal of Language and Linguistic Studies*, *11*(1), 75-90. https://files.eric.ed.gov/fulltext/EJ1105203.pdf

Hutchinson, A. (2016, July 12). *Facebook conducts research into the opportunities of mobile gaming* [infographic]. Social Media Today. https://www.socialmediatoday.com/social-business/facebook-conducts-research-opportunities-mobile-gaming-infographic

Hung, C. Y., Kuo, F. O., Sun, J. C. Y., & Yu, P. T. (2013). An interactive game approach for improving students' learning performance in multi-touch game-based learning. *IEEE Transactions on Learning Technologies*, *7*(1), 31-37. IEEE. https://doi.org/10.1109/TLT.2013.2294806

Joseph, S. R., & Uther, M. (2009). Mobile devices for language learning: Multimedia approaches. *Research and Practice in Technology Enhanced Learning*, *4*(1), 7-32. World Scientific. https://doi.org/10.1142/S179320680900060X

Kaliszewski, M. (2019, October 14). *Statistics of drug use in high school*. American Addiction Centers. https://americanaddictioncenters.org/blog/statistics-of-drug-use-in-high-school

Kang, S. H., Pashler, H., Cepeda, N. J., Rohrer, D., Carpenter, S. K., & Mozer, M. C. (2011). Does incorrect guessing impair fact learning? *Journal of Educational Psychology*, *103*(1), 48. American Psychological Association. https://doi.org/10.1037/a0021977

Keep Talking and Nobody Explodes [Video Game]. (2015). Steel Crate Games.

Kirriemuir, J., & McFarlane, A. (2004). *Literature review in games and learning*. Futurelab. https://telearn.archives-ouvertes.fr/hal-00190453

Klepek, P. (2017). *How a major JRPG wound up getting totally re-written months after release*. VICE. https://www.vice.com/en/article/7xwzmb/how-a-major-jrpg-wound-up-getting-totally-re-written-months-after-release

Klimova, B., & Kacet, J. (2017). Efficacy of computer games on language learning. *Turkish Online Journal of Educational Technology-TOJET*, *16*(4), 19-26. https://files.eric.ed.gov/fulltext/EJ1160637.pdf

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, *22*(6), 529-540. Wiley. http://dx.doi.org/10.1111/j.1944-9720.1989.tb02780.x

Kongmee, I., Strachan, R., Pickard, A., & Montgomery, C. (2012). A case study of using online communities and virtual environment in massively multiplayer role playing games (MMORPGs) as a learning and teaching tool for second language learners. *International Journal of Virtual and Personal Learning Environments (IJVPLE), 3*(4), 1-15. IGI Global. http://doi.org/10.4018/jvple.2012100101

Kopel, D. B. (1993). Japanese gun control. *Asia Pacific Law Review*, *2*(2), 26-52. Routledge. https://doi.org/10.1080/18758444.1993.11787979

Koptur, E. (2016). *Transferable learning skills of an MMORPG: A World of Warcraft quest* [Doctoral Dissertation, Kent State University).

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Addison-Wesley Longman Limited.

Lai, C., Ni, R., & Zhao, Y. (2012). Digital games and language learning. In M. Thomas, H. Reinders, M. Warschauer (Eds.), *Contemporary computer-assisted language learning* (pp. 183-200). Bloomsbury.

Lansdale, M., & Baguley, T. (2008). Dilution as a model of long-term forgetting. *Psychological Review*, *115*(4), 864-892. American Psychological Association. https://doi.org/10.1037/a0013325

Lantolf, J. P. (1994). Sociocultural theory and second language learning: Introduction to the special issue. *The Modern Language Journal*, *78*(4), 418-420. Wiley. http://www.jstor.org/stable/328580

Laufer, B. (2000). Task effect on instructed vocabulary learning: The hypothesis of 'involvement'. In AILA'99 Tokyo Organizing Committee (Eds.) *Selected Papers from AILA'99 Tokyo* (pp. 47-62). Waseda University Press.

Leone, M. (2017, January 9). *Final fantasy 7: An oral history*. Polygon. https://www.polygon.com/a/final-fantasy-7

Li, J. (2019). A Path Analysis of Players' Perception of Using Massively Multiplayer Role-Playing Games (MMORPGs) to Learn Second Language (L2) English Vocabulary. In K. Graziano (Ed.), *Society for information technology & teacher education international conference* (pp. 1931-1937). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/207910/

Li, J. (2019). *The role of massively multiplayer role-playing games in facilitating vocabulary acquisition for English language learners: A mixed-methods study*. [Doctoral dissertation, Clemson University). https://tigerprints.clemson.edu/all_dissertations/2342

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta--analysis. *Language Learning*, *60*(2), 309-365. Wiley. https://doi.org/10.1111/j.1467-9922.2010.00561.x

Life Is Strange [Video Game]. (2015). Dontnod Entertainment, Decknine, Square Enix.

Ling, R. (2005). The sociolinguistics of SMS: An analysis of SMS use by a random sample of Norwegians. In R. Ling & P.E. Pedersen (Eds.), *Mobile communications* (pp. 335-349). Springer. https://doi.org/10.1007/1-84628-248-9_22

Little, J., & Ligon Bjork, E. (2012). The persisting benefits of using multiple-choice tests as learning events. In *Proceedings of the annual meeting of the cognitive science society, 34* (pp. 683-688). Univeristy of California. https://escholarship.org/uc/item/50d9x93k

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492-527. American Psychological Association. https://doi.org/10.1037/0033-295X.95.4.492

Lohse, K., Shirzad, N., Verster, A., Hodges, N., & Van der Loos, H. M. (2013). Video games and rehabilitation: using design principles to enhance engagement in physical therapy. *Journal of Neurologic Physical Therapy*, *37*(4), 166-175. Wolters Kluwer. https://doi.org/10.1097/NPT.0000000000000017

Long, M. H. (1980). *Input, interaction, and second language acquisition*. University of California, Los Angeles.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*(3), 251-285. Cambridge University Press. https://doi.org/10.1017/S0272263100009165

Long, M. H. (2016). In defense of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics*, *36*, 5-33. Cambridge University Press. https://doi.org/10.1017/S0267190515000057

Machado, A., Santos, P., & Dias, J. (2017). On the structure of role playing game quests. *Revista de Ciências da Computação*, *12*, 81-98. Universidade Aberta. http://hdl.handle.net/10400.2/11090

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545-562. Wiley. https://doi.org/10.2307/330224

Maiier, N. (2017). The methodological training of future university teachers to implement intercultural foreign language education: the competency-based approach. *Advanced Education*, *4*(8), 4-9. https://doi.org/10.20535/2410-8286.101358

Marchand, A., & Hennig-Thurau, T. (2013). Value creation in the video game industry: Industry economics, consumer benefits, and research opportunities. *Journal of Interactive Marketing*, *27*(3), 141-157. Elsevier. https://doi.org/10.1016/j.intmar.2013.05.001

Mattiello, E. (2005). The pervasiveness of slang in standard and non-standard English. *Mots Palabras Words, 6*, 7-41.

Mayer, R. E. (2002). Multimedia learning. In B.H. Ross (Ed.), *Psychology of learning and motivation, 41* (pp. 85-139). Academic Press. https://doi.org/10.1016/S0079-7421(02)80005-6

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E., Lee, H., & Peebles, A. (2014). Multimedia learning in a second language: A cognitive load perspective. *Applied Cognitive Psychology*, *28*(5), 653-660. Wiley. https://doi.org/10.1002/acp.3050

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, *105*(4), 724-760. American Psychological Association. https://doi.org/10.1037/0033-295X.105.4.734-760

Mcleod, S. (2019, July 10). *Effect size*. Simply Psychology. https://www.simplypsychology.org/effect-size.html

Meeter, M., & Murre, J. M. (2004). Consolidation of long-term memory: Evidence and alternatives. *Psychological Bulletin*, *130*(6), 843-857. American Psychological Association. https://doi.org/10.1037/0033-2909.130.6.843

Merriam-Webster [Dictionary]. (2021). https://www.merriam-webster.com

Miller, M., & Hegelheimer, V. (2006). The SIMs meet ESL Incorporating authentic computer simulation games into the language classroom. *Interactive Technology and Smart Education, 3*(4), 311-328. Emerald Publishing. http://dx.doi.org/10.1108/17415650680000070

Mobile Operating System Market Share Worldwide. (2021). StatCounter Global Stats. https://gs.statcounter.com/os-market-share/mobile/worldwide

Moghaddas, B., & Bashirnezhad, H. (2016). The pedagogical applications of mobile-assisted language learning (MALL) in improving the Iranian EFL learners' oral performance. *International Journal of Applied Linguistics and Translation*, *2*(1), 8-14. Science Publishing Group. https://doi.org/10.11648/j.ijalt.20160201.12

Monfared, A., Cervantes, S. E., Lee, S. M., & Jackson, M. (2018). Establishing a praxis between SLA theory and CALL-based practices. In J. Perren, K. Kelch, J. Byun, S.

Cervantes, & S. Safavi (Eds.), *Applications of CALL theory in ESL and EFL environments* (pp. 59-79). IGI Global. https://doi.org/10.4018/978-1-5225-2933-0.ch004

Montola, M. (2007). Tangible pleasures of pervasive role-playing. In *DiGRA '07 - Proceedings of the 2007 DiGRA international conference: Situated Play* (pp. 178-185). http://www.digra.org/wp-content/uploads/digital-library/07312.38125.pdf

Moore, P. J. (2018). Task-based language teaching (TBLT). In J.I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1-7). Wiley. https://doi.org/10.1002/9781118784235.eelt0175

Murdock, B. B. (1985). The contributions of Hermann Ebbinghaus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(3), 469-471. American Psychological Association. https://doi.org/10.1037/0278-7393.11.3.469

Murdock, B., Smith, D., & Bai, J. (2001). Judgments of frequency and recency in a distributed memory model. *Journal of Mathematical Psychology*, *45*(4), 564-602. Elsevier. https://doi.org/10.1006/jmps.2000.1339

Nagle, S. J., & Sanders, S. L. (1986). Comprehension theory and second language pedagogy. *TESOL Quarterly*, *20*(1), 9-26. TESOL. https://doi.org/10.2307/3586386

Nah, K. C., White, P., & Sussex, R. (2008). The potential of using a mobile phone to access the internet for learning EFL listening skills within a Korean context. *ReCALL: Journal of Eurocall*, *20*(3), 331-347. Cambridge University Press. https://doi.org/10.1017/S0958344008000633

Newon, L. (2011). Multimodal creativity and identities of expertise in the digital ecology of a World of Warcraft guild. In C. Thurlow & K. Mroczek (Eds.), *Digital discourse: Language in the new media* (pp. 131-153). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199795437.003.0007

Nichol, S., & Blashki, K. (2008). Enhancing electronic learning for generation Y games geeks. In N. Kock (Ed.), *Encyclopedia of e-collaboration* (pp. 246-252). IGI Global. http://hdl.handle.net/10536/DRO/DU:30017051

Nietfeld, J. L. (2017). The role of self-regulated learning in digital games. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 271-284). Routledge.

Nikolov, M., & Csapó, B. (2018). The relationships between 8th graders' L1 and L2 reading skills, inductive reasoning and socio-economic status in early English and German as a foreign language programs. *System, 73*, 48-57. Elsevier. https://doi.org/10.1016/j.system.2017.11.001

Nishimura, Y. (2015). Style, creativity and play. In A. Georgakopoulou & T. Spilioti (Eds.), *The Routledge handbook of language and digital communication* (pp. 117-131). Routledge. https://doi.org/10.4324/9781315694344

North, D. (2015, April 14). *155m Americans play video games, and 80% of households own a gaming device*. VentureBeat. https://venturebeat.com/2015/04/14/155-million-americans-play-video-games-and-4-out-of-5-households-own-a-gaming-device/

Nunan, D. (1988). Principles for designing language teaching materials. *Guidelines: A Periodical for Classroom Language Teachers*, *10*(2), 1-24. https://eric.ed.gov/?id=EJ468940

One Night Ultimate Werewolf [Board Game]. (2013). Bézier Games.

Pearl, L., & Goldwater, S. (2011). Statistical learning, inductive bias, and Bayesian inference in language acquisition. In J.L. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford handbook of developmental linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199601264.013.28

Peterson, M. (2010). Massively multiplayer online role-playing games as arenas for second language learning. *Computer Assisted Language Learning*, *23*(5), 429-439. Routledge. https://doi.org/10.1080/09588221.2010.520673

Peterson, M. (2010). Task-based language teaching in network-based CALL: An analysis of research on learner interaction in synchronous CMC. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 41-62). Bloomsbury. http://dx.doi.org/10.5040/9781474212366.ch-003

Peterson, M. (2012). Language learner interaction in a massively multiplayer online role-playing game. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 70-92). Palgrave Macmillan. https://doi.org/10.1057/9781137005267_5

Peterson, M. (2012). Learner interaction in a massively multiplayer online role playing game (MMORPG): A sociocultural discourse analysis. *ReCALL*, *24*(3), 361-380. Cambridge University Press. https://doi.org/10.1017/S0958344012000195

Peterson, M. (2013). Computer games and language learning: Theoretical rationales. In *Computer Games and Language Learning* (pp. 51-60). Palgrave Macmillan. https://doi.org/10.1057/9781137005175_4

Peterson, M. (2016). The use of massively multiplayer online role-playing games in CALL: An analysis of research. *Computer Assisted Language Learning*, *29*(7), 1181-1194. Routledge. https://doi.org/10.1080/09588221.2016.1197949

Peterson, M., White, J., Mirzaei, M. S., & Wang, Q. (2020). A review of research on the application of digital games in foreign language education. In M. Kruk & M. Peterson (Eds.), *New technological applications for foreign and second language learning and teaching* (pp. 69-92). IGI Global. https://doi.org/10.4018/978-1-7998-2591-3.ch004

Phillips, T. (2015, March 20). *Life Is Strange dev talks lip-syncing, release schedule and publisher demands*. Eurogamer. https://www.eurogamer.net/articles/2015-03-18-publishers-wanted-life-is-stranges-main-character-to-be-a-guy-qtes-added

Phillips, T. (2016, April 18). *Life Is Strange team talk fan theories, that season finale and more*. Eurogamer. https://www.eurogamer.net/articles/2016-04-09-life-is-strange

Pica, T. (1987). Second-language acquisition, social interaction, and the classroom. *Applied Linguistics*, *8*(1), 3-21. Oxford University Press. https://doi.org/10.1093/applin/8.1.3

Piechurska-Kuciel, E. (2011). Willingness to communicate in L2 and self-perceived levels of FL skills in Polish adolescents. In J. Aranslo & A. Wojtaszek (Eds.), *Aspects of culture in second language acquisition and foreign language learning* (pp. 235-250). Springer. https://doi.org/10.1007/978-3-642-20201-8_1

Pietschmann, D., Liebold, B., & Valtin, G. (2017). The MMORPG designer's journey: casualization and its consequences for social interactions. In R. Kowert & T. Quandt (Eds.), *New perspectives on the social aspects of digital gaming* (pp. 82-96). Routledge. https://doi.org/10.4324/9781315629308

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*(1), 1-28. University of Hawai‘i at Mānoa. https://files.eric.ed.gov/fulltext/EJ759833.pdf

Piirainen–Marsh, A., & Tainio, L. (2009). Collaborative game-play as a site for participation and situated learning of a Second Language. *Scandinavian Journal of Educational Research*, *53*(2), 167-183. Routledge. https://doi.org/10.1080/00313830902757584

Plester, B., Wood, C., & Joshi, P. (2009). Exploring the relationship between children's knowledge of text message abbreviations and school literacy outcomes. *British Journal of Developmental Psychology*, *27*(1), 145-161. The British Psychological Society. https://doi.org/10.1348/026151008x320507

Portnow, J. (2011). Gaming languages and language games. In M. Adams (Ed.), *From Elvish to Klingon: Exploring invented languages* (pp. 135-159). Oxford University Press.

Ranalli, J. (2008). Learning English with The Sims: exploiting authentic computer simulation games for L2 learning. *Computer Assisted Language Learning*, *21*(5), 441-455. Routledge. https://doi.org/10.1080/09588220802447859

Rankin, Y. A., Gold, R., & Gooch, B. (2006). 3D role-playing games as language learning tools. In E. Gröller & L. Szirmay-Kalos (Eds.), *Eurographics (Education Papers)*, *25*(3) (pp. 33-38). The Eurographics Association. http://dx.doi.org/10.2312/eged.20061005

Rankin, Y. A., McNeal, M., Shute, M. W., & Gooch, B. (2008, August). User centered game design: evaluating massive multiplayer online role playing games for second language acquisition. In D. Schwarz, K. Schrier, C. Swain, & M. Wagner (Eds.), *Proceedings of the 2008 ACM SIGGRAPH symposium on video games* (pp. 43-49). Association for Computing Machinery. https://doi.org/10.1145/1401843.1401851

Rankin, Y. A., Morrison, D., McNeal, M., Gooch, B., & Shute, M. W. (2009, April). Time will tell: In-game social interactions that facilitate second language acquisition. In J. Whitehead & R. M. Young (Eds.), *Proceedings of the 4th international conference on foundations of digital games* (pp. 161-168). Association for Computing Machinery. http://dx.doi.org/10.1145/1536513.1536546

Rankin, Y. A., & Shute, M. W. (2010). Re-purposing a recreational video game as a serious game for second language acquisition. In J. Cannon-Bowers & C. Bowers (Eds.), *Serious game design and development: Technologies for training and learning* (pp. 178-195). IGI Global. https:/doi.org/10.4018/978-1-61520-739-8.ch010

Reinders, H., & Wattana, S. (2012). Talk to me! Games and students' willingness to communicate. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 156-188). Palgrave Macmillan. http://dx.doi.org/10.1057/9781137005267_9

Reinders, H., & Wattana, S. (2014). Can I say something? The effects of digital game play on willingness to communicate. *Language Learning & Technology, 18*(2), 101-123. https://hdl.handle.net/10652/2962

Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, *27*(1), 38-57. Cambridge University Press. http://dx.doi.org/10.1017/S0958344014000226

Reinhardt, J. (2017). Digital gaming in L2 teaching and learning. In C.A. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 202-216). Wiley. https://doi.org/10.1002/9781118914069

Reinhardt, J., & Thorne, S. (2016). Metaphors for digital games and language learning. In F. Farr & L. Murray (Eds.), *The Routledge handbook of language learning and technology* (pp. 441-456). Routledge. https://doi.org/10.4324/9781315657899

Robert, R. (2012). *Video games and their very real economic wars - Paris innovation review*. RSS. http://parisinnovationreview.com/articles-en/video-games-and-their-very-real-economic-wars

Rosell-Aguilar, F. (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO Journal*, *34*(2), 243-258. Equinox Publishing. https://doi.org/10.1558/cj.27623

Roskos, K. A., Christie, J. F., Widman, S., & Holding, A. (2010). Three decades in: Priming for meta-analysis in play-literacy research. *Journal of Early Childhood Literacy*, *10*(1), 55-96. American Psychological Association. https://psycnet.apa.org/doi/10.1177/1468798409357580

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595-598. American Psychological Association. https://psycnet.apa.org/doi/10.1038/33396

Saltzman, M. (2021). *E3 2021: Video games are bigger business than ever, topping movies and music combined*. USA Today. https://www.usatoday.com/videos/tech/2021/06/10/e-3-2021-video-games-big-business-topping-film-and-music-combined/7637695002/

Schneider, J. (2013). Remembrance of things past: A history of the Socratic method in the United States. *Curriculum Inquiry*, *43*(5), 613-640. Routledge. https://doi.org/10.1111/curi.12030

Scovel, T. (1994). The role of culture in second language pedagogy. *System*, *22*(2), 205-219.

Shaffer, D. W., & Gee, J. P. (2006). *How computer games help children learn*. Palgrave Macmillan.

Sharples, M., Taylor, J., & Vavoula, G. (2005, October). Towards a theory of mobile learning. In H. van der Merwe, & T. Brown (Eds.), *Mobile technology: The future of learning in your hands* (pp. 58).

Shimane, T., & Wada, K. (2007). Substance use among night high school students in Japan. *Japanese Journal of Alcohol Studies & Drug Dependence*, *42*(3), 152-164.

Sjöblom, B., & Aronsson, K. (2012). Participant categorisations of gaming competence: Noob and imba as learner identities. In O. Erstad & J. Sefton-Green (Eds.), *Identity, community, and learning lives in the digital Age* (pp. 181-196). Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139026239.013

Squire, K. (2006). From content to context: Video games as designed experience. *Educational Researcher*, *35*(8), 19-29. Sage Publishing. https://doi.org/10.3102%2F0013189X035008019

Steinkuehler, C. A. (2004). Learning in massively multiplayer online games. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the sixth international conference of the learning sciences* (pp. 521–528). Association for Computing Machinery.

Stipek, D., Feiler, R., Daniels, D., & Milburn, S. (1995). Effects of different instructional approaches on young children's achievement and motivation. *Child Development*, *66*(1), 209-223. Wiley. https://doi.org/10.2307/1131201

Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of platform. *Language Learning & Technology*, *14*(2), 95-110. http://llt.msu.edu/vol14num2/stockwell.pdf

Stockwell, G., & Hubbard, P. (2013). *Some emerging principles for mobile-assisted language learning*. The International Research Foundation for English Language Education. http://www.tirfonline.org/english-in-the-workforce/mobile-assisted-language-learning/

Strong, S. A. (2018). *Gamer-generated language and the localisation of massively multiplayer online role-playing games* [Doctoral Dissertation, University College London]. https://discovery.ucl.ac.uk/id/eprint/10046921

Suh, S., Kim, S. W., & Kim, N. J. (2010). Effectiveness of MMORPG-based instruction in elementary English education in Korea. *Journal of Computer Assisted Learning*, *26*(5), 370-378. Wiley. https://doi.org/10.1111/j.1365-2729.2010.00353.x

Susi, T., Johannesson, M., & Backlund, P. (2007). *Serious games: An overview*. https://www.diva-portal.org/smash/get/diva2:2416/fulltext01.pdf

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). Cambridge University Press. https://psycnet.apa.org/doi/10.1017/CBO9780511816819.003

Sweller, J. (2011). Cognitive load theory. In K.D. Federmeier & L. Sahakyan (Eds.) *Psychology of learning and motivation* (pp. 37-76). Elsevier. https://psycnet.apa.org/doi/10.1016/B978-0-12-387691-1.00002-8

Sykes, J. M. (2014). TBLT and synthetic immersive environments: What can in-game task restarts tell us about design and implementation?. In M. González & L. Ortega (Eds.), *Technology-mediated TBLT* (pp. 149-182). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.6.06syk

Sykes, J. M., & Dubreil, S. (2019). Pragmatics learning in digital games and virtual environments. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 387-399). Routledge. https://doi.org/10.4324/9781351164085

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, *24*(3), 302-321.Cambridge University Press. https://doi.org/10.1017/S095834401200016X

Tagliamonte, S. A., & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen language. *American Speech*, *83*(1), 3-34. Duke University Press. https://doi.org/10.1215/00031283-2008-001

Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, involvement, and flow in digital games. In R. Bernhaupt (Ed.), *Evaluating user experience in games* (pp. 23-46). Springer. https://doi.org/10.1007/978-1-84882-963-3_3

Tane, K., & Yamamoto, Y. (2012, February 3). 東亜プラン特集 － 元・東亜プラン 開発者インタビュー： 弓削雅稔. *Shooting Gameside, 4*, 41–48.

Tangle Tower [Video Game]. (2019). SFB Games.

Tarng, P. Y., Chen, K. T., & Huang, P. (2008, October). An analysis of WoW players' game hours. In M. Claypool (Ed.), *Proceedings of the 7th ACM SIGCOMM workshop on network and system support for games* (pp. 47-52). Association for Computing Machinery. https://doi.org/10.1145/1517494.1517504

Taylor, D. (1994). Inauthentic authenticity or authentic inauthenticity. *TESL-EJ*, *1*(2), 1-11.

Tekken [Video Game]. (1994). Bandai Namco Studios.

Tetris [Video Game]. (1984). Alexey Pajitnov, Vladimir Pokhilko.

Theodore, S. (2016, October 31). *Why have video game budgets skyrocketed in recent years?* Forbes. https://www.forbes.com/sites/quora/2016/10/31/why-have-video-game-budgets-skyrocketed-in-recent-years/?sh=6b1e18923ea5

The Sims [Video Game]. (2000). Maxis.

Thomas, M. (2012). Contextualizing digital game-based language learning: Transformational paradigm shift or business as usual?. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 11-31). Palgrave Macmillan. https://doi.org/10.1057/9781137005267_2

Thomes, T. P. (2015). In-house publishing and competition in the video game industry. *Information Economics and Policy*, *32*(C), 46-57. Elsevier. https://doi.org/10.1016/j.infoecopol.2015.07.005

Thompson, C. G., & von Gillern, S. (2020). Video-game based instruction for vocabulary acquisition with English language learners: A Bayesian meta-analysis. *Educational Research Review, 30*(2): 100332. Elsevier. http://dx.doi.org/10.1016/j.edurev.2020.100332

Thorne, S. L., & Watters, E. (2013). Review of language at play: Digital games in second and foreign language teaching and learning. *Language Learning & Technology, 17*(3), 47-51. University of Hawaii National Foreign Language Resource Center. http://hdl.handle.net/10125/44338

Turkay, S., Hoffman, D., Kinzer, C. K., Chantes, P., & Vicari, C. (2014). Toward understanding the potential of games for learning: Learning theory, game design characteristics, and situating video games in classrooms. *Computers in the Schools*, *31*(1-2), 2-22. Routledge. http://dx.doi.org/10.1080/07380569.2014.890879

Tüzün, H. (2007). Blending video games with learning: Issues and challenges with classroom implementations in the Turkish context. *British Journal of Educational Technology*, *38*(3), 465- 477. British Educational Research Association. http://dx.doi.org/10.1111/j.1467-8535.2007.00710.x

Tychsen, A., Hitchens, M., & Brolund, T. (2008). Motivations for play in computer role-playing games. In B. Kapralos, M. Katchabaw, & J. Rajnovich (Eds.), *Proceedings of the 2008 conference on future play: Research, play, share* (pp. 57-64). Association for Computing Machinery. https://doi.org/10.1145/1496984.1496995

Tyndale, E., & Ramsoomair, F. (2016). Keys to successful interactive storytelling: A study of the booming “choose-your-own-adventure” video game industry. *i-manager’s Journal of Educational Technology*, *13*(3), 28-34. https://files.eric.ed.gov/fulltext/EJ1133237.pdf

Unsworth, S. (2007). Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch. *Language Acquisition*, *14*(2), 215-217. Routledge. https://doi.org/10.1080/10489220701353891

Van Avermaet, P., Colpin, M., Van Gorp, K., Bogaert, N., & Van den Branden, K. (2006). The role of the teacher in task-based language teaching. In K. van de Branden (Ed.), *Task-based language education: From theory to practice* (pp. 175-196). Cambridge University Press. http://hdl.handle.net/1854/LU-4284282

Vegel, A. (2018). Critical perspective on language learning: TBLT and digital games. In S. Sasayama (Ed.), *Proceedings of the TBLT in Asia 2018 Conference* (pp. 113-124).

Video game industry statistics, trends and data in 2021. (2021, November 17). WePC. https://www.wepc.com/news/video-game-statistics/

Warnick, B. R., Kim, S. H., & Robinson, S. (2015). Gun violence and the meaning of American schools. *Educational Theory*, *65*(4), 371-386. Wiley. https://doi.org/10.1111/edth.12122

Weda, S. (2018). The effects of students' motivation and family socioeconomic status on English academic achievement. *The Asian EFL Journal*, *20*(12.3), 132-145. http://eprints.unm.ac.id/id/eprint/14860

Weisberg, D. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education*, *7*(2), 104-112. International Mind, Brain, and Education Society. https://doi.org/10.1111/mbe.12015

Weisberg, D. S., Zosh, J. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Talking it up: play, language development, and the role of adult support. *American Journal of Pla*y, *6*(1), 39-54. The Strong. https://eric.ed.gov/?id=EJ1016058

White, M. (2019, January 30). *How big should the control group be in a randomized field experiment?* Mark H. White II, PhD. https://www.markhw.com/blog/control-size

Wichadee, S., & Pattanapichet, F. (2018). Enhancement of performance and motivation through application of digital games in an English language class. *Teaching English with Technolog*y, *18*(1), 77-92. IATEFL Poland Computer Special Interest Group. https://eric.ed.gov/?id=EJ1170635

Widdowson, H. G. (1998). Context, community, and authentic language. *TESOL Quarterly*, *32*(4), 705-716. TESOL. https://doi.org/10.2307/3588001

World of Warcraft [Video Game]. (2004). Blizzard Entertainment.

Wrobetz, K. R. (2018). On the use of graphic novels in the ESL classroom. *The Journal of the College of Foreign Languages, Himeji Dokkyo University, 31,* 109-131. http://id.nii.ac.jp/1567/00000083/

Wrobetz, K.R. (2021). Board games versus smartphone applications: Gaming media in EFL contexts. In P. Clements, R. Derrah, & P. Ferguson (Eds.), *Communities of teachers, & learners*. *JALT*. https://doi.org/10.37546/JALTPCP2020-43

Wrobetz, K.R. (in press). Gaming in a foreign language: Observations on second-language vocabulary processing in single-player RPGs. In B.L. Reynolds (Ed.), *Vocabulary learning in the wild*. Springer.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning, & Technology*, *13*(2), 48-67. University of Hawaii National Foreign Language Resource Center. https://eric.ed.gov/?id=EJ851895

Yan, V. X., Yu, Y., Garcia, M. A., & Bjork, R. A. (2014). Why does guessing incorrectly enhance, rather than impair, retention?. *Memory & Cognition*, *42*(8), 1373-1383. Springer. https://psycnet.apa.org/doi/10.3758/s13421-014-0454-6

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, *86*(1), 54-66. Wiley. http://dx.doi.org/10.1111/1540-4781.00136

Zagal, J. P., & Altizer, R. (2014, April). Examining 'RPG elements': Systems of character progression. In T. Barnes & I. Bogost (Eds.), *Proceedings of the 9th International Conference on the Foundations of Digital Games*. Society for the Advancement of the Science of Digital Games. http://www.fdg2014.org/papers/fdg2014_paper_38.pdf

Zero Wing [Video Game]. (1989). Toaplan.

Zhao, Y., & Lai, C. (2009). MMORPGs and foreign language education. In R.G. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (pp. 402-421). IGI Global. https://doi.org/10.4018/978-1-59904-808-6.ch024

Zheng, D., Bischoff, M., & Gilliland, B. (2015). Vocabulary learning in massively multiplayer online games: Context and action before words. *Educational Technology Research and Development*, *63*(5), 771-790. Springer. http://dx.doi.org/10.1007/s11423-015-9387-4

Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going?. *Computer Assisted Language Learning*, *34*(5-6), 751-777. Routledge. https://doi.org/10.1080/09588221.2019.1640745

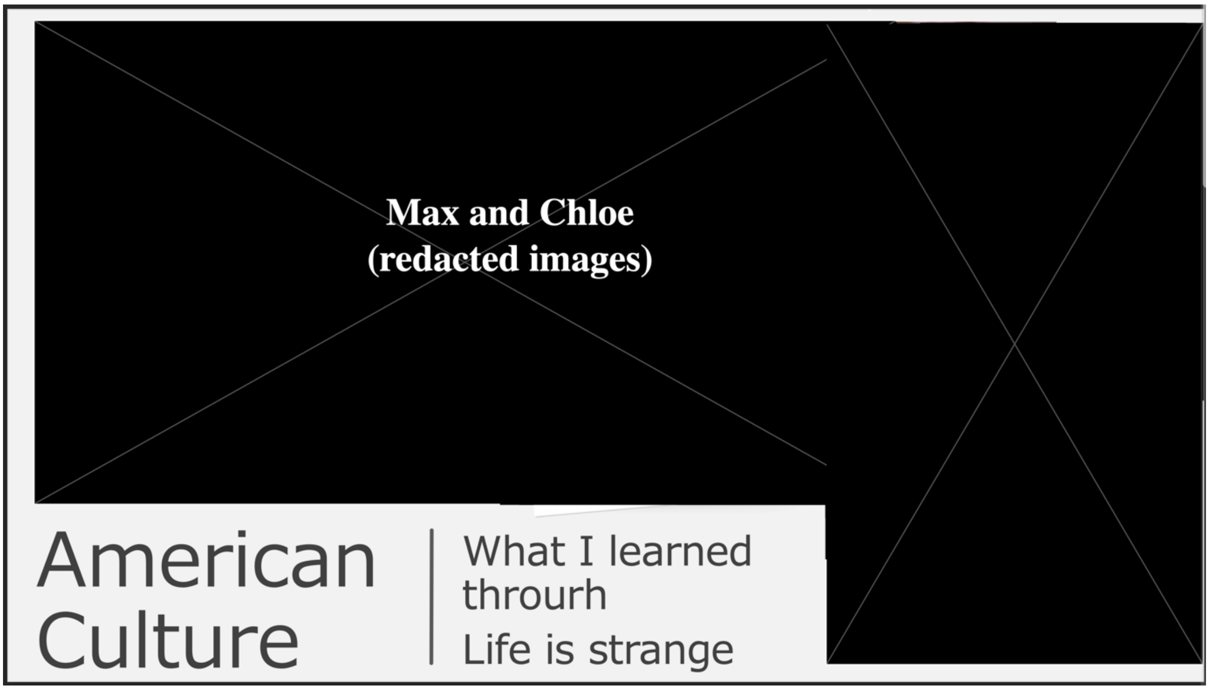
Appendices

Appendix A Publication consent statement

Publication consent statement (Japanese)	Publication consent statement (English)
このテストの目的は、京都大学で行われている博士課程の研究のためのデータを収集することです。すべてのデータは完全に匿名で管理されます。このテストを受けることで、あなたのデータが上記の研究に使用されることに同意するものとします。	The purpose of this test is to gather data for doctoral research being conducted at Kyoto University. All data will be handled completely anonymously. By taking this test, you hereby give your consent for the use of your data in the above-mentioned research.

Note. The English translation is presented for clarification purposes only.

Appendix B LIS oral presentation from PS group A



2 what was your impression of US culture after playing?

- drive to school
- guns
- Bullying (same as in Japan)
- call friends name many time in conversation
- unique irony

3

3 How did your impression of US culture change after wards?

- There is no culture of wearing uniforms.
- They do not clean the classroom by them selves.
- They are very close to the teachers.
- They can feely put on hair color , clothe and decoration.

4

4

What are the merits of
using video games to portray culture?

1 have a feel for gaming
------Be easy to remember because of the game
------Be interested in languages

2 Be able to learn intuitively
------Be intuitive to the culture
------Make a clear difference from Japan

Start-up menu (redacted image)

Language settings, English subtitles option (redacted image)

5

5 What are the demerit of using video games?

1 eyes are going to go bad
----My eyes are easy to get tired.

2 Depends on the games
----If you get addicted to games, you can do it a long time and lead to dependence.

Max and Chloe (redacted image)

Mobile gaming console (redacted image)

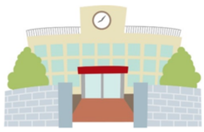
6

Appendix C LIS oral presentation from PS group B

Image of American culture

Before of playing the game.

I had the image of being able to get driver's license from 18 years old.



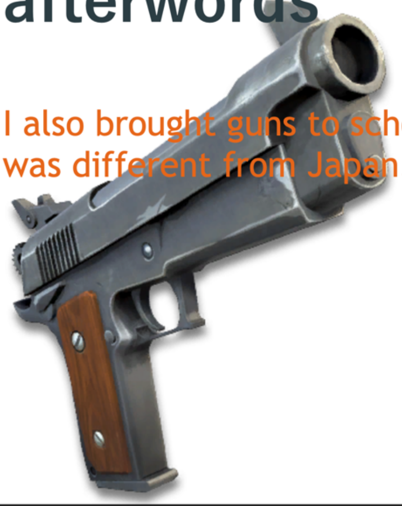
Max and Chloe in truck (redacted image)

After playing the game.

Can get a driver's license from 16 years old.

⇒American high school students can go to school by car.

US culture change afterwords



I also brought guns to school, did drugs, I thought it was different from Japan.

Merits of using video games to portray culture.

- It is easy to leave in memory if you see in the picture.
- Pictures are easier and easier to understand than letters.
- You feel like you actually experienced it.

Max playing guitar
(redacted image)

Demerits of using video games

Adversely affects Japanese culture.

Max and Chloe
(redacted images)

Appendix D Vocabulary pre-test, post-test, and OSD post-test

	Vocabulary test items	Correct response	Distractor responses		
1	mad skills	上手 (skillful)	怒っている (angry)	裏技 (secret trick)	連勝中 (winning streak)
2	blacked out	気を失った (lost consciousness)	怒った (got angry)	黒くなった (got black)	外に行った (went outside)
3	fetish	フェチ (fetish)	癖 (bad habit)	防犯行動 (crime prevention activities)	おべっかを使う (suck up to)
4	dude	お前 (you)	変人 (oddball)	おいおい、やつめ (hey chump)	男 (man)
5	got hella cash	超お金持ち (exceedingly rich)	現金 (cash)	貧乏 (poor)	宝物 (treasure)
6	timeout	休憩 (break)	気を失う (lose consciousness)	緊急 (emergency)	時間 (time)
7	I got this	任せろ (leave it to me)	パニックになる (freak out)	早くいけ (go fast)	上手 (skillful)
8	on fire	連勝中 (winning streak)	火 (fire)	大ファン (big fan)	賢い (clever)
9	poof	急に消えた (disappeared suddenly)	帰った (went home)	煙 (smoke)	みっけ (there it is)
10	selfie	自撮り (selfie)	写メ (photograph)	自分で (by one's self)	マジか (seriously?)
11	messed up	正気とは思えない (insane)	ストレスが溜まっている (stressed out)	ぐちゃぐちゃ (messy)	怖い (scary)
12	psycho	頭おかしい人 (crazy person)	おいおい、やつめ (hey chump)	強敵 (formidable opponent)	臆病 (cowardice)
13	nosey	おせっかいな (nosey)	好奇心 (curious)	鼻 (nose)	みっけ (there it is)
14	mess around	浮気する (cheat on)	外出する (go out)	ふざける (screw around)	遊ぶ (play)
15	kickstart	促す (encourage)	始める (start)	おべっかを使う (suck up to)	蹴り出す (kick out)
16	your silence is deafening	しーん (deathly quiet)	教科書 (textbook)	持ち込み (bring-in items)	甘ったるい (sentimental)
17	bizarro	奇妙 (strange)	促す (encourage)	変人 (oddball)	講義 (lecture)
18	boatload	大量にある (in abundance)	船 (boat)	正直に (honestly)	率直に (candidly)
19	booty	宝物 (treasure)	お金 (money)	ケツ (rump)	おせっかいな (nosey)
20	everything's cool	大丈夫 (okay)	涼しい (cool temperature)	任せろ (leave it to me)	早く行け (go fast)
21	photo op	シャッターチャンス (photo opportunity)	カメラ (camera)	おべっかを使う (suck up to)	モデル (model)
22	screwing around	ふざける (screw around)	誠に (honestly)	冗談 (joke)	ストレスが溜まっている (stressed out)
23	blaze	大麻を吸う (smoke marijuana)	ぼーっとする (stare off into space)	大音量で音量を聴く (listen to music at high volume)	炎 (flame)
24	mushy	甘ったるい (sentimental)	フェチ (fetish)	柔らかい (soft)	必須 (required)
25	chill	リラックスしながら時間を潰す (kill time while relaxing)	待ち合わせる (meet up with)	冷やす (cool off)	マジか (seriously?)
26	boney ass	ガリガリの体 (skinny body)	お尻 (backside)	連勝中 (winning streak)	骨 (bone)

27	meltdown	号泣 (bawling)	崩壊 (collapse)	正気とは思えない (insane)	悩み (distress)
28	move your ass	早くいけ (go fast)	ふざける (screw around)	運動しろ (go exercise)	お前ら (you guys)
29	freak out	パニックになる (freak out)	落ち着け (calm down)	変人 (oddball)	爆発する (explode)
30	upfront	率直に (candidly)	大ファン (big fan)	フロント (front)	優しい (kind)
31	boohoo	泣き虫 (cry baby)	頭おかしい人 (crazy person)	文句 (complaint)	なめるな (don't patronize me)
32	pissed	怒る (get angry)	絶やす (wipe out)	小便 (urine)	タバコ (tobacco)
33	kissing ass	おべっかを使う (suck up to)	キスする (to kiss)	怒る (get mad)	いちゃいちゃする (flirt)
34	sucked hard	大変だった (was difficult)	盛り上がる (get excited)	大雨 (heavy rain)	ストレスが溜まっている (stressed out)
35	wheels	車 (car)	車輪 (wheel/s)	自転車 (bicycle)	マジか (seriously?)
36	cereal	マジか (seriously?)	穀物 (grain)	フェチ (fetish)	朝食 (breakfast)
37	wiped out	一掃される (be eraticated)	吹く (blow)	ふざける (screw around)	火災 (fire)
38	lame	ダサい (tacky)	カッコい (so cool)	偉そうに (pompous)	おせっかいな (nosey)
39	groupies	大ファン (big fan)	扱き使う (work someone hard)	生徒 (student)	シャッターチャンス (photo opportunity)
40	stressed out	ストレスが溜まっている (stressed out)	笑止 (ridiculous)	大変だった (was difficult)	疲れている (tired)
41	badass	カッコい (so cool)	パニックになる (freak out)	変わっている (strange)	ダメ (no good)
42	freak	変人 (oddball)	パニックになる (freak out)	ボロボロ (ragged)	やってみろ (just try it)
43	lighthouse	灯台 (lighthouse)	悩み (distress)	光っている家 (lit up house)	塔 (tower)
44	blasting	大音量で音楽を聴く (listen to music at high volume)	大麻を吸う (smoke marijuana)	聞かせて (let someone listen)	爆発させる (detonate)
45	boom	みっけ (there it is)	やっぱり (as expected)	爆発 (explosion)	急に消えた (disappeared suddenly)

Appendix E Pre-game survey

Test items		English translation	Question type	Responses available
1	スマホゲームをよくやっています。	I often play smartphone games.	10-point Likert scale	1=strongly disagree, 10=strongly agree
2	スマホゲームをする時、どの言語でやっていますか?	When you play smartphone games, what language do you play in?	multiple-choice with optional input	Japanese, English, I don't play smartphone games, other (optional input)
3	スマホゲームをする時、どこでやっていますか?	When you play smartphone games, where do you play?	open-ended response	
4	スマホゲームを英語でする経験はありますか?	Have you had the experience of playing smartphone games in English?	multiple-choice	yes, no
5	前の質問で「あります」と答えた方、どのスマホゲームを英語でやったことありますか?	For those who answered "yes" in the previous question, what smartphone games have you played in English?	open-ended response	
6	スマホゲームを英語でするだけで英語力が伸びる。	Playing smartphone games in English is enough to improve English proficiency.	10-point Likert scale	1=strongly disagree, 10=strongly agree
7	スマホゲームを英語でると、どの語学力が伸びると思いますか?	When playing smartphone games in English, what linguistic skills do you think will be improved?	multiple-response multiple-choice with optional input	vocabulary, listening, speaking, reading, writing, other (optional input)
8	アメリカ文化に関する知識を下に記述してください。	Describe your knowledge of US culture below.	open-ended response	
9	アメリカの教育制度に関する知識を下に記述してください。	Describe your knowledge of the US educational system below.	open-ended response	
10	アメリカ人の学生に対する印象を下に記述してください。	Describe your impressions of US students below.	open-ended response	
11	上記のアメリカについての情報源は何ですか?	What informational sources do you use to inform your opinions of the above questions about the US?	open-ended response	
12	株式会社スクウエア・エニックスの「ライフ イズ ストレンジ」というゲームをやったことありますか?	Have you ever played the game "Life Is Strange" published by Square Enix Co., Ltd.?	multiple-choice	yes, no

13	前の質問で「あります」と答えた方、どの言語でゲームをやりましたか?	For those who answered “yes” in the previous question, what language did you play the game in?	multiple-choice with optional input	Japanese, English, other (optional input)
14	スマホゲームは、英語力向上のために活用できると思いますか? その意見を具体的に説明してください。	Do you think that smartphone games can be used to improve your English skills? Please explain your opinion in detail.	open-ended response	
15	英語力向上のためにスマホゲームを利用することの賛否両論は何ですか?	What are the advantages and disadvantages of using smartphone games to improve your English skills?	open-ended response	
16	一般的な教室での指導と比較し、英語力向上のために利用されるスマホゲームはより効果的だと思いますか? その意見を具体的に説明してください。	Compared to general classroom instruction, do you think that smartphone games used to improve English skills are more effective? Please explain your opinion in detail.	open-ended response	

Appendix F Immediate vocabulary test 1

Vocabulary test item	Correct response	Distractor responses		
1 on fire	連勝中 (winning streak)	火 (fire)	大ファン (big fan)	賢い (clever)
2 stressed out	ストレスが溜まっている (stressed out)	笑止 (ridiculous)	大変だった (was difficult)	疲れている (tired)
3 meltdown	号泣 (bawling)	崩壊 (collapse)	正気とは思えない (insane)	悩み (distress)
4 psycho	頭おかしい人 (crazy person)	おいおい、やつめ (hey chump)	強敵 (formidable opponent)	臆病 (cowardice)
5 kickstart	促す (encourage)	始める (start)	おべっかを使う (suck up to)	蹴り出す (kick out)
6 lighthouse	灯台 (lighthouse)	悩み (distress)	光っている家 (lit up house)	塔 (tower)
7 timeout	休憩 (break)	気を失う (lose consciousness)	緊急 (emergency)	時間 (time)
8 boohoo	泣き虫 (cry baby)	頭おかしい人 (crazy person)	文句 (complaint)	なめるな (don't patronize me)
9 upfront	率直に (candidly)	大ファン (big fan)	フロント (front)	優しい (kind)
10 got hella cash	超お金持ち (exceedingly rich)	現金 (cash)	貧乏 (poor)	宝物 (treasure)
11 freak	変人 (oddball)	パニックになる (freak out)	ボロボロ (ragged)	やってみろ (just try it)
12 selfie	自撮り (selfie)	写メ (photograph)	自分で (by one's self)	マジか (seriously?)
13 your silence is deafening	しーん (deathly quiet)	教科書 (textbook)	持ち込み (bring-in items)	甘ったるい (sentimental)
14 freak out	パニックになる (freak out)	落ち着け (calm down)	変人 (oddball)	爆発する (explode)
15 photo op	シャッターチャンス (photo opportunity)	カメラ (camera)	おべっかを使う (suck up to)	モデル (model)
16 kissing ass	おべっかを使う (suck up to)	キスする (to kiss)	怒る (get mad)	いちゃいちゃする (flirt)
17 everything's cool	大丈夫 (okay)	涼しい (cool temperature)	任せろ (leave it to me)	早く行け (go fast)

Appendix G Immediate vocabulary test 2

Vocabulary test item	Correct response	Distractor responses		
1 wheels	車 (car)	車輪 (wheel/s)	自転車 (bicycle)	マジか (seriously?)
2 groupies	大ファン (big fan)	扱き使う (work someone hard)	生徒 (student)	シャッターチャンス (photo opportunity)
3 dude	お前 (you)	変人 (oddball)	おいおい、やつめ (hey chump)	男 (man)
4 booty	宝物 (treasure)	お金 (money)	ケツ (rump)	おせっかいな (nosey)
5 bizarro	奇妙 (strange)	促す (encourage)	変人 (oddball)	講義 (lecture)
6 I got this	任せろ (leave it to me)	パニックになる (freak out)	早くいけ (go fast)	上手 (skillful)
7 pissed	怒る (get angry)	絶やす (wipe out)	小便 (urine)	タバコ (tobacco)
8 boney ass	ガリガリの体 (skinny body)	お尻 (backside)	連勝中 (winning streak)	骨 (bone)
9 move your ass	早くいけ (go fast)	ふざける (screw around)	運動しろ (go exercise)	お前ら (you guys)
10 boom	みっけ (there it is)	やっぱり (as expected)	爆発 (explosion)	急に消えた (disappeared suddenly)
11 chill	リラックスしながら時間を潰す (kill time while relaxing)	待ち合わせる (meet up with)	冷やす (cool off)	マジか (seriously?)
12 mess around	浮気する (cheat on)	外出する (go out)	ふざける (screw around)	遊ぶ (play)
13 lame	ダサい (tacky)	カッコい (so cool)	偉そうに (pompous)	おせっかいな (nosey)

Appendix H Immediate vocabulary test 3

Vocabulary test item	Correct response	Distractor responses		
1 messed up	正気とは思えない (insane)	ストレスが溜まっている (stressed out)	ぐちゃぐちゃ (messy)	怖い (scary)
2 blacked out	気を失った (lost consciousness)	怒った (got angry)	黒くなった (got black)	外に行った (went outside)
3 nosey	おせっかいな (nosey)	好奇心 (curious)	鼻 (nose)	みっけ (there it is)
4 wiped out	一掃される (be eradicated)	吹く (blow)	ふざける (screw around)	火災 (fire)
5 mushy	甘ったるい (sentimental)	フェチ (fetish)	柔らかい (soft)	必須 (required)
6 mad skills	上手 (skillful)	怒っている (angry)	裏技 (secret trick)	連勝中 (winning streak)
7 screwing around	ふざける (screw around)	誠に (honestly)	冗談 (joke)	ストレスが溜まっている (stressed out)
8 fetish	フェチ (fetish)	癖 (bad habit)	防犯行動 (crime prevention activities)	おべっかを使う (suck up to)
9 blasting	大音量で音楽を聴く (listen to music at high volume)	大麻を吸う (smoke marijuana)	聞かせて (let someone listen)	爆発させる (detonate)
10 poof	急に消えた (disappeared suddenly)	帰った (went home)	煙 (smoke)	みっけ (there it is)
11 boatload	大量にある (in abundance)	船 (boat)	正直に (honestly)	率直に (candidly)
12 badass	カッコい (so cool)	パニックになる (freak out)	変わっている (strange)	ダメ (no good)
13 blaze	大麻を吸う (smoke marijuana)	ぼーっとする (stare off into space)	大音量で音量を聴く (listen to music at high volume)	炎 (flame)
14 cereal	マジか (seriously?)	穀物 (grain)	フェチ (fetish)	朝食 (breakfast)
15 sucked hard	大変だった (was difficult)	盛り上がる (get excited)	大雨 (heavy rain)	ストレスが溜まっている (stressed out)

Appendix I Comprehension test A

Character identification test items		
Image of Max	Image of Mr. Jefferson	Image of Victoria
Item 1: What is the name of this character? Correct response: Max Distractor responses: Mr. Jefferson, Victoria, Kate, Nathan, Mr. Madsen	Item 2: What is the name of this character? Correct response: Mr. Jefferson Distractor responses: Max, Victoria, Kate, Nathan, Mr. Madsen	Item 3: What is the name of this character? Correct response: Victoria Distractor responses: Max, Mr. Jefferson, Kate, Nathan, Mr. Madsen
Image of Kate	Image of Nathan	Image of Mr. Madsen
Item 4: What is the name of this character? Correct response: Kate Distractor responses: Mr. Jefferson, Victoria, Max, Nathan, Mr. Madsen	Item 5: What is the name of this character? Correct response: Nathan Distractor responses: Max, Victoria, Kate, Mr. Jefferson, Mr. Madsen	Item 6: What is the name of this character? Correct response: Mr. Madsen Distractor responses: Max, Mr. Jefferson, Kate, Nathan, Victoria

Narrative content comprehension questions			
	Test items	Correct response	Distractor responses
7	Which character is Max's "rival"?	Victoria	Max, Mr. Jefferson, Kate, Nathan, Mr. Madsen
8	Which character is being teased in class?	Kate	Mr. Jefferson, Victoria, Max, Nathan, Mr. Madsen
9	What subject does Mr. Jefferson teach?	photography	math, history, English
10	What level is the school?	high school	middle school, university, prep school
11	What is the name of the school?	Blackwell Academy	Highschool, Louis Daguerre, Prescott High School
12	What is Max's "special power"?	She can reverse time.	She is psychic. / Her dreams become reality. / She is great at photography.
13	What is the name of the photo contest Mr. Jefferson recommends to his class?	Everyday Heroes	John Lennon Life Contest, Blackwell Academy, Selfies
14	In what city is the photo contest held in?	San Francisco	New York City, Los Angeles, Seattle
15	Why was a girl shot in the bathroom?	There was an argument over drugs and money.	The boy is a psycho. / The girl is the boy's ex-girlfriend. / The boy wanted answers to a test.

Appendix J Comprehension test B

Character identification test items		
Image of Juliet	Image of Dana	Image of Samuel
Item 1: What is the name of this character?	Item 2: What is the name of this character?	Item 3: What is the name of this character?
Correct response: Juliet	Correct response: Dana	Correct response: Samuel
Distractor responses: Dana, Samuel, Zachary, Warren, Chloe	Distractor responses: Juliet, Samuel, Zachary, Warren, Chloe	Distractor responses: Dana, Juliet, Zachary, Warren, Chloe
Image of Zachary	Image of Warren	Image of Chloe
Item 4: What is the name of this character?	Item 5: What is the name of this character?	Item 6: What is the name of this character?
Correct response: Zachary	Correct response: Warren	Correct response: Chloe
Distractor responses: Dana, Samuel, Juliet, Warren, Chloe	Distractor responses: Dana, Samuel, Zachary, Juliet, Chloe	Distractor responses: Dana, Samuel, Zachary, Warren, Juliet

Narrative content comprehension questions			
	Test items	Correct response	Distractor responses
7	Who is Victoria's ex-boyfriend?	Zachary	Dana, Samuel, Juliet, Warren, Chloe
8	Who does the flash drive belong to?	Warren	Dana, Samuel, Zachary, Juliet, Chloe
9	Why did Juliet lock Dana in the dormitory room?	Juliet thought that Dana was cheating with Zachary.	Juliet thought that Dana stole some of her money. / Juliet was angry that Dana is friends with Victoria. / Juliet was angry that Dana is pregnant.
10	Why does Juliet forgive Dana?	Max found proof that Dana is innocent.	Victoria admitted she was behind the "sexting." / Zachary told Juliet that Dana is innocent. / Kate had proof that Dana is innocent.
11	What year is Warren's car?	1978	2000, 1985, 1999
12	Where does Warren want to go with Max in his car?	the movies	a restaurant, Seattle, New Castle
13	What kind of content was on the flash drive?	movies	music, pictures, emails
14	Why is Nathan angry with Max?	He knows that Max was in the bathroom earlier.	He doesn't like the teacher Mr. Jefferson. / He wants to get even with Chloe. / He doesn't like Max's photographs.
15	What movie does Warren recommend Max to watch?	Cannibal Holocaust	Akira, Twilight Zone, Evangelion

Appendix K Comprehension test C

Character identification test items	
Image of Chloe	Image of David
Item 1: What is the name of this character? Correct response: Chloe Distractor responses: Max, David, Principal Wells, Mr. Jefferson, Alyssa	Item 2: What is the name of this character? Correct response: David Distractor responses: Max, Chloe, Principal Wells, Mr. Jefferson, Alyssa

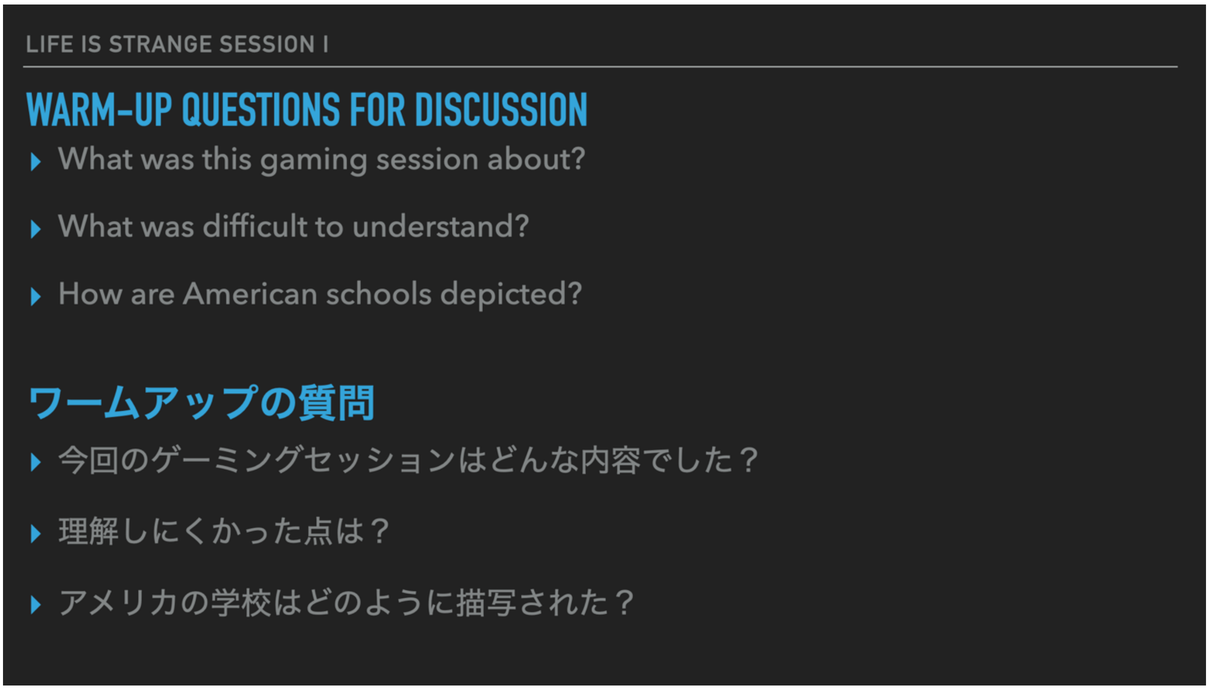
Narrative content comprehension questions			
	Test items	Correct response	Distractor responses
3	What is Nathan's family name?	Prescott	Caulfield, Price, Amber
4	What US city did Max move to before returning to the city of Arcadia Bay?	Seattle	San Francisco, Chicago, Oregon
5	What is the relationship between Chloe and David?	David is Chloe's stepfather.	David is Chloe's father. / David is Chloe's boss. / David is Chloe's boyfriend.
6	What is the name of Chloe's friend who has disappeared?	Rachel	Juliet, Max, Nathan
7	In what room did Max find the precision screw drivers?	garage	living room, Chloe's bedroom, kitchen
8	Before becoming the security guard at Blackwell Academy, what was David Madsen's job?	He was in the military.	He was a model. / He was in a punk band. / He was unemployed.
9	Who did the camera that Chloe gave Max originally belong to?	Chloe's real father	Rachel, David Madsen, Mr. Jefferson
10	How old was Chloe when Max moved away?	14	16, 12, 20
11	What city was Rachel planning on moving to before she disappeared?	Los Angeles	Paris, Arcadia Bay, Seattle
12	When did Rachel disappear?	6 months ago	14 years ago, 1 year ago, yesterday
13	Why did Chloe contact Nathan?	She wanted to steal money from him.	She wanted to sell him drugs. / She wanted to go on a date with him. / She wanted him to lend her money.
14	When does Max predict that Arcadia Bay will be destroyed by a tornado?	October 11th	October 31st, November 13th, Friday the 13th
15	What US state does the game take place in?	Oregon	Washington, California, Seattle

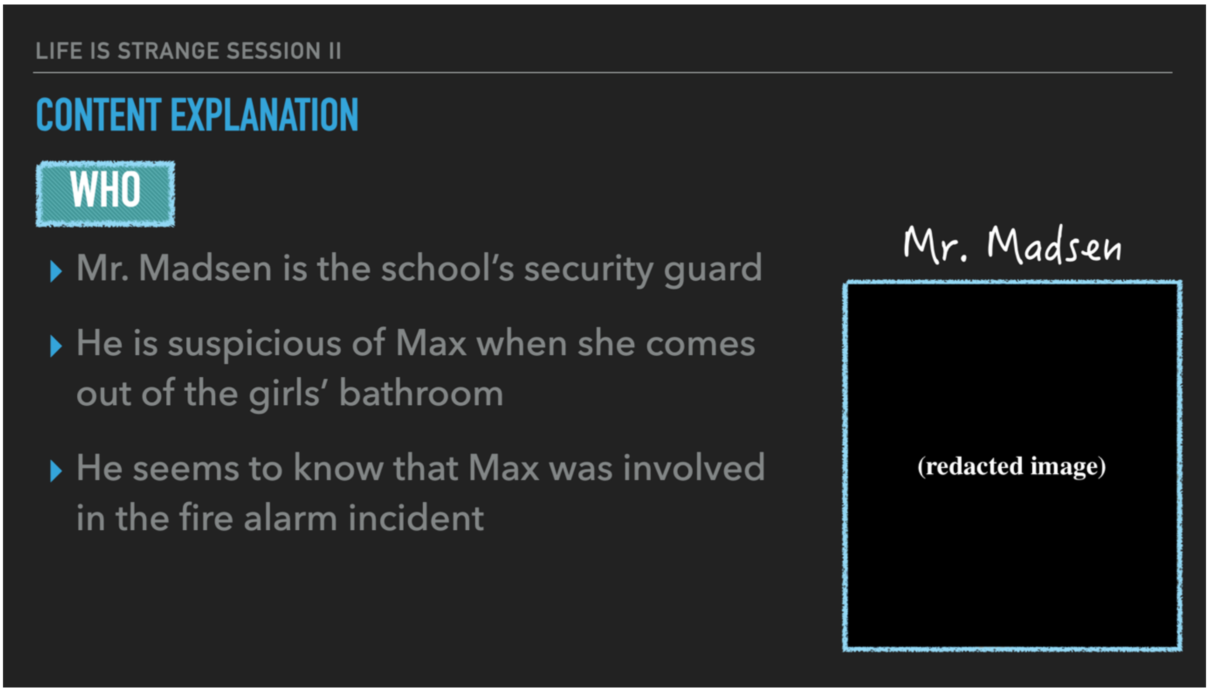
Appendix L Post-game survey

	Test items	English translation	Question type	Responses available
1	このゲームを最後までしたいと思っていますか?	Do you feel that you want to play this game until the end?	multiple-choice	yes, no
2	ゲーミングセッション1をどこでやりましたか（例えば、電車やお家など）?	Where did you play gaming session 1 (on the train, at home, etc.)?	open-ended response	
3	ゲーミングセッション2をどこでやりましたか（例えば、電車やお家など）?	Where did you play gaming session 2 (on the train, at home, etc.)?	open-ended response	
4	ゲーミングセッション3をどこでやりましたか（例えば、電車やお家など）?	Where did you play gaming session 3 (on the train, at home, etc.)?	open-ended response	
5	このゲームの内容をよく理解できました。	I comprehended the content of the game well.	10-point Likert scale	1=strongly disagree, 10=strongly agree
6	このゲームに関して、特に理解しにくいことは何だと思いますか?	Concerning the game, what parts were particularly hard to comprehend?	open-ended response	
7	これからスマホゲームをもっとしようと思っています。	I'm planning on playing more smartphone games in the future.	10-point Likert scale	1=strongly disagree, 10=strongly agree
8	これからスマホゲームをする時、どの言語でするつもりですか?	Which language do you plan to use when playing smartphone games in the future?	multiple-choice with optional input	Japanese, English, other (optional input)
9	スマホゲームを英語でするだけで英語力が伸びる。	Playing smartphone games in English is enough to improve English proficiency.	10-point Likert scale	1=strongly disagree, 10=strongly agree
10	このゲームで使われた英語は難しかったです。	The English used in this game was difficult.	10-point Likert scale	1=strongly disagree, 10=strongly agree
11	スマホゲームを英語でする と、どの語学力が伸びると思いますか?	When playing smartphone games in English, what linguistic skills do you think will be improved?	multiple-response multiple-choice with optional input	vocabulary, listening, speaking, reading, writing, other (optional input)
12	アメリカ文化に関する知識を下に記述してください。	Describe your knowledge of US culture below.	open-ended response	

13	アメリカの教育制度に関する知識を下に記述してください。	Describe your knowledge of the US educational system below.	open-ended response	
14	アメリカ人の学生に対する印象を下に記述してください。	Describe your impressions of US students below.	open-ended response	
15	上記のアメリカについての情報源は何ですか?	What informational sources do you use to inform your opinions of the above questions about the US?	open-ended response	
16	ゲームの言語設定のおかげで実感した、英語力の伸びを下のスケールで示して下さい。	Please indicate with the scale below how much you believe your English proficiency has increased due to the game's language setting.	10-point Likert scale	1=little improvement, 10=substantial improvement
17	スマホゲームは、英語力向上のために活用できると思いますか? その意見を具体的に説明してください。	Do you think that smartphone games can be used to improve your English skills? Please explain your opinion in detail.	open-ended response	
18	英語力向上のためにスマホゲームを利用することの賛否両論は何ですか?	What are the advantages and disadvantages of using smartphone games to improve your English skills?	open-ended response	
19	一般的な教室での指導と比較し、英語力向上のために利用されるスマホゲームはより効果的だと思いますか? その意見を具体的に説明してください。	Compared to general classroom instruction, do you think that smartphone games used to improve English skills are more effective? Please explain your opinion in detail.	open-ended response	

Appendix M Excerpts from in-class discussion presentations





LIFE IS STRANGE SESSION II

CONTENT EXPLANATION

WHAT

- The story so far has introduced that the main character, Max, has the power to reverse time
- It is unclear why Max has this power, but she uses it to save the girl who was shot in the bathroom

Max
(redacted image)

Time Manipulation

LIFE IS STRANGE SESSION III

CONTENT EXPLANATION

WHERE

- Blackwell Academy is what is known as a "boarding school"
- Like a university, students at Blackwell Academy have the option of staying in the Prescott Dormitory
- Students also drive to school

Blackwell Academy

School building
(redacted image)

Parking lot
(redacted image)

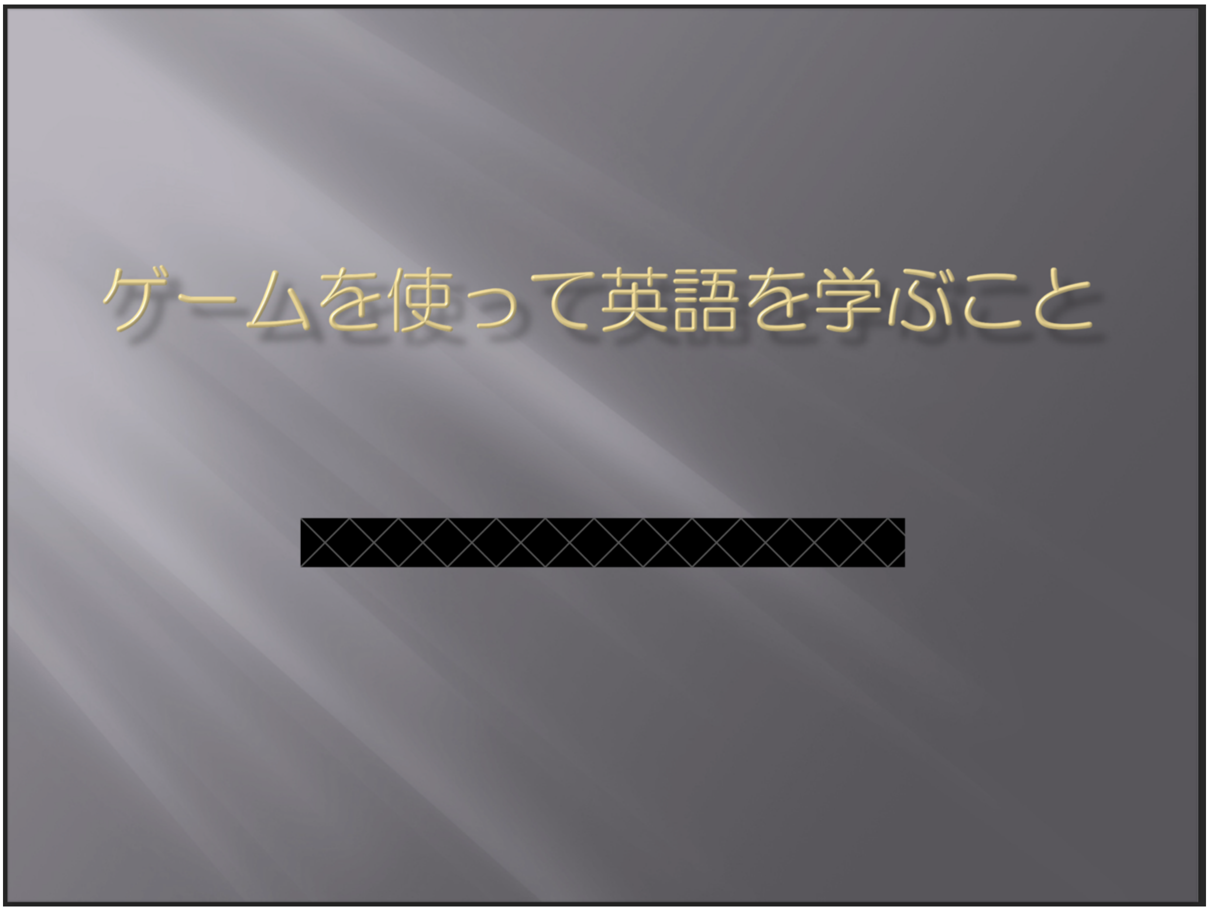
Appendix N Oral presentation topics and questions for PS group

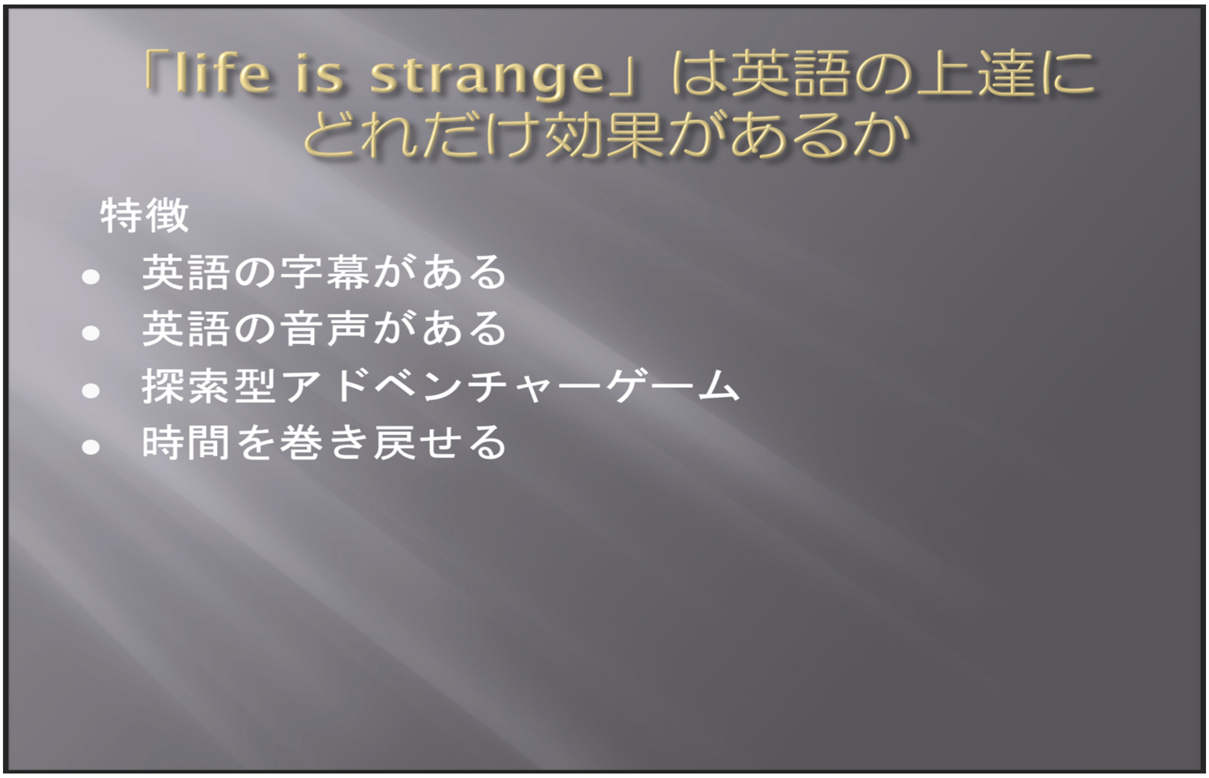
Presentation topic 1: News article review	
Presentation instructions	Please choose one English language news article from a credible news agency (e.g., CNN, BBC, the New York Times, NHK, etc.)
Presentation objectives	Your presentation should review the most important elements of the article and explain the article's content in a manner that is easy to understand for everyone in class.
Presentation questions	1 Who is this article about? 2 What is this article about? 3 When does this article take place? 4 Where does this article take place? 5 Why was this article written? (What is the objective?)

Presentation topic 2: Advertisement	
Presentation instructions	Please choose a product or service to make an advertisement for. Your advertisement presentation should be entirely in English.
Presentation objectives	Your presentation should advertise the most important elements of your chosen product/service and explain why the product/service is important in a manner that is easy to understand for everyone in class.
Presentation questions	1 What is the product/service? (a simple introduction is okay) 2 What is the demographic for the product/service? 3 What are the merits of the product/service? 4 Why is this product/service superior to other products? 5 Why does the customer need this product/service?

Presentation topic 3: US culture in video games	
Presentation instructions	Please use the game *Life Is Strange* (Square Enix, 2015) to discuss how US culture is portrayed in video games.
Presentation objectives	Your presentation should review how US culture is portrayed in the game *Life Is Strange*. Furthermore, your presentation should talk about the merits/demerits of using video games to portray culture.
Presentation questions	1 What was your impression of US culture before playing? 2 What was your impression of US culture after playing? 3 How did your impression of US culture change afterwards? 4 What are the merits of using video games to portray culture? 5 What are the demerits of using video games?

Appendix O LIS oral presentation from LG group A





1,今学期にしたゲーム

- 楽しかったゲーム「Life is Strange」
- 英語の上達に役立ったゲーム「Keep Talking and Nobody Explodes」

3,ゲームと従来の教材を比較

ゲームの方が良い学習ができる

理由

- 関心を持つ生徒が増える
- 挫折しにくい

似ているところ

- 文章を読む
- 単語を覚える

違うところ

- 自分で考えなければならない

4,英語を学ぶためにゲームを使う

考えが変わった

理由

ゲームは遊びだから従来の学習より英語が身につかないのではないかと思っていたが、ゲームは単語などを覚えていないとゲームに参加できないため、従来の学習より熱心に単語を覚えようとするし、ゲームを通して学んだ単語をアウトプットできるためゲームの方が英語が身に付くという考えに変わった。

英語を学ぶためにゲームを使う可能性は高くなると思う

7,どんな科目がゲームは役立つのか

社会を学習するときにゲームは役立つ

理由

- 社会は暗記科目のため普通に覚えるよりもゲームにした方が飽きずに楽しく学習できる

Appendix P LIS oral presentation from LG group B

Learn English using games

[redacted]

What I felt when I played Life is strange

- Good for introduction to learning English

- Increases listening ability

US culture

- Enter the house with shoes
- Drive to high school
- Classes with a small number of people

1

- The most fun game is Life is strange.
- I think Life is strange improves my English skills.
  →Because many English expressions come out.

4

・I changed my opinion of using games to learn English.
→Because it is fun to learn English.

・I think you will be more likely to learn English using games.

6

- I think that games will be used more often in the future to learn English in educational settings.

Because you can enjoy learning without thinking deeply about difficult English

8

- I had played a game in English before.
  I had played FPS game.
  And I think it helped me improve my English