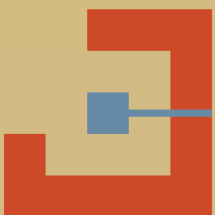




Exploring the Intersection of Games and Technology in Language Education



Selected papers from the JALTCALL2022 Conference

ISBN 978-4-901352-70-3

Playful CALL

Selected papers from the JALTCALL2022 Conference, Japan [online]



JALT CALL is a Special Interest Group (SIG) of the Japan Association for Language Teaching (JALT) that focuses on Computer-Assisted Language Learning (CALL) and technology in language learning.

<http://jaltcall.org/>



JALT is dedicated to the improvement of language teaching and learning.

JALT Central Office

Urban Edge Bldg. 5F, 1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan

<http://jalt.org/>

© 2023 by the JALT CALL SIG and individual authors.

First published in Japan in 2023 by the JALT CALL SIG.

Layout by Paul Mason

Preface

Play. Who plays? What, why, and how do we play? What can computer-assisted language learning teachers and researchers learn from play? These core questions inspired the theme of the 2022 JALTCALL conference. But why this theme? Why now?

In this “Ludic” century, games and play are becoming the most pervasive and arguably, the most important media in our lives. But what are games? A simple definition may be a “designed experience” or a set of rules which generate interesting decisions in players. We can also think of games as systems that users manipulate toward certain gratifying outcomes. Extrapolating this thought towards CALL research, then, CALL is inextricably linked to the study of systems (both in terms of computers and technology, but also in terms of languages as systems) and their effect upon one another. As an example, research may explore how a computer system may be utilized towards improving learners’ pronunciation, reading comprehension, or pragmatic awareness. However, as Zimmerman (2013) writes: “It is not enough to merely be a systems-literate person; to understand systems in an analytic sense. We also must learn to be playful in them. A playful system is a human system, a social system rife with contradictions and with possibility.” Thus, the JALTCALL 2022 conference theme was born – to promote language teachers and researchers to think about research and teaching from the concept of play and its inherent implications, with the understanding that playfulness and technology can bring together creativity, innovation, and meaningfulness.

The papers presented at the 2022 conference and those published here are intended to expand the dialogue around the intersections of CALL, play, and the term “ludic” as we continue headlong into the Ludic Century. Virtual reality and augmented reality represent cutting-edge technology explored in CALL research, and in this volume of papers, the reader will see a playful approach to their implementation towards language learning gains. Similarly, papers explored playful approaches to extracurricular learning, where one innovative study looked at the formation of a community of practice around the multiplayer game named *Minecraft*.

We would like to thank the authors who worked hard to contribute to this volume. We are deeply grateful for their efforts. We also thank the reviewers who generously donated their time and expertise to the process of evaluating the manuscripts. We were also blessed to have our associate editors, who were keen to play with us in the *game* of publishing this volume, and who worked hard with the authors to improve their manuscripts for publication. We are indebted to the program committee and all the attendees who made this conference and this volume possible. Finally, we thank the JALTCALL Executive Board, who fully supported the conference and the publication.

We look forward to seeing you all at the JALTCALL conference in 2023.

Editors-in-Chief
Todd Dwayne Cooper, University of Toyama
James York, Meiji University

Editors-in-Chief

Todd Dwayne Cooper is an associate professor at the University of Toyama. Todd explores the usage of technology to assess and improve nonverbal communication. His latest *kakenhi* funded research focuses on analyzing facial expression.

James York is a senior assistant professor at Meiji University. He conducts research on ludic approaches to language teaching. He is also Editor-in-Chief of Ludic Language Pedagogy, an open-access journal that publishes research on the integration of games and play in language teaching contexts.

Associate Editors

Araceli Salas is a professor-researcher at the Benemérita Universidad Autónoma de Puebla, México. She specializes in discourse analysis and teacher education. She is part of the Editorial Review Board of the MEXTESOL Journal and the Editor in Chief of Lenguas en Contexto published by her university (<http://www.facultadlenguas.com/lencontexto/>).

Anthony Brian Gallagher MAODE, PGCE, PGCODE, BSc(Hons) is a specially appointed assistant professor in the Faculty of Foreign Studies at Meijo University, and a Sports Instructor at Kinjo Gakuin University, Nagoya. His interdisciplinary research interests include: The user experience & course quality assurance; Assistive, disruptive, & cosmetic technologies; Physiology & Sports Science, and Performance in Education.

Dara Tafazoli (PhD) is Research Officer on the Virtual Reality (VR) School Study, Australia. He is currently working on developing and implementing the CALL literacy framework for language teachers at the School of Education, the University of Newcastle, Australia. His research interests are Computer-Assisted Language Learning (CALL) and CALL teacher education and professional development. Further info: www.daratafazoli.com

Elena Alcalde Peñalver is an Associate Professor at the University of Alcalá and holds a PhD on Financial Translation from the University of Granada (Spain). She also works as a freelance translator and interpreter and has teaching and research experience at national and international level in the fields of foreign languages and translation.

George MacLean is a professor at the University of the Ryukyus. He has taught at primary-junior high and university levels in Japan and in the international school system. His research interests include SLA and ICT implementation. He regularly presents at national and international language teaching and technology events.

Işıl Günseli Kaçar is a lecturer, teacher educator and faculty member at the Department of Foreign Language Education at Middle East Technical University (METU) in Turkey. She is interested in pre-service language teacher education, pre-service teacher identity, English as a Lingua Franca, World Englishes, intercultural communication, flipped learning, e-mentoring, technology-mediated language learning/teaching, virtual exchange, and academic writing. She worked as a tutor at the academic writing center at METU previously. She is currently coordinating a number of national and international research projects on pre-service teacher education and e-mentoring.

Joe Tomei is a professor in the Faculty of British and American Studies at Kumamoto Gakuen University. He has taught EFL in France, Spain, Korea and Japan at the primary, secondary, and tertiary levels. In addition to his interest in computer-mediated communication, he also is interested in the application of functional/typological grammar to language teaching, practical activities in the language classroom, and writing instruction.

Michael Blodgett is a lecturer at Nara Medical University. His research interests include motivation and authentic communication activities enhanced by technology. Currently, he is exploring the use of podcasts in language education.

Satchie Haga is a full time lecturer at Rikkyo University. She conducts cross cultural research that examines sensemaking, knowledge construction, and intersectionality in distributed sociotechnical systems.

Layout

Paul Mason, Production Editor, *The JALTCALL Journal*

Technical Editor (website programming and design)

Gary Ross, JALTCALL SIG Webmaster

Contents

1	Interactivity's effect on playing versus watching a real-time strategy game: A preliminary experiment <i>Shawn Andersson, Ritsumeikan University</i>	1
2	Using online forums to promote Japanese university student motivation to use English outside the classroom <i>Jack Pudelek, Kwansei Gakuin University</i>	26
3	Self-access learning and Minecraft: Observations and student perceptions <i>Robert Remmerswaal and Robert Dykes</i>	43
4	Ready, set...? End-user testing of an AR treasure hunt learning game <i>Samuel Taylor and Adam Stone</i>	68
5	Evaluation of a VR language learning environment: Effect of feedback on learners' flow state <i>James York, Meiji University; Hirotaka Takuma, Takuya Yoshino and Hiroshi Nakayama, Tokyo Denki University</i>	85

1

Interactivity's effect on playing versus watching a real-time strategy game: A preliminary experiment

Shawn Andersson, Ritsumeikan University

Abstract

Digital Game-based Language Learning (DGBLL) involves playing digital games to learn a language, utilizing game elements inductive to motivation and learning. The field has been gaining attention recently, but applications traditionally only involved users playing the games themselves. Recently, watching gameplay online has become a global sensation, yet its effectiveness for foreign language pedagogy remains unexplored. One significant difference between playing and watching a game is the inclusion or exclusion of physical interactivity, which involves utilizing a controller to manipulate the game. Studies have previously addressed interactivity in games, but predominantly utilize serious games made for educational purposes. This study includes a preliminary experiment testing the effects of physical interactivity of a commercial off-the-shelf (COTS) game where participants were assigned player or watcher roles. Eleven ($n = 11$) participants took part in the study, and a mixed-method approach was utilized for data collection consisting of a vocabulary test, questionnaire, interviews, and researcher observations. The results indicate a slightly higher vocabulary achievement for the watchers, while the players reported greater mental effort toward their task relative to learning English and rated their task at a higher difficulty. Conversely, players expressed better post-treatment attitudes and perceptions of playing games to learn English. Nevertheless, the interviews indicated that having low-skilled players negatively affected the experience of the watchers, and the players felt too busy controlling the game to learn English. Finally, recommendations were made for future follow-up studies.

Digital Game-based Language Learning (DGBLL) では、デジタルゲームをプレイすることで、ゲームの要素を利用して言語学習のモチベーションと学習を促進する。これは昨今注目されている分野だが、

従来の研究は、ユーザーが自らゲームをプレイするというものであった。一方で最近、オンラインで他者がゲームをプレイしているのを視聴する、いわゆるゲーム実況中継動画の視聴が世界で流行している。しかし、ゲーム視聴と外国語教育におけるその有効性はまだ調査されていない。ゲームをプレイすることと視聴することの大きな違いの1つは、コントローラーを使用してゲームを操作することを含む、物理的な双方向性を含めるか除外するかである。また、これまでの研究ではゲームのインタラクティブ性が研究されてきたが、それには主に教育目的で作成されたserious gamesが利用されている。

本研究では、ゲームの視聴が与える英語学習の影響に着目し、commercial off-the-shelf (COTS) gameをプレイ、または視聴した際に英語学習者に与える物理的な相互作用の影響を調査した。11名 (n = 11) の参加者を、5人のプレイヤーと6人のウォッチャーに分け、語彙テスト、アンケート、インタビュー、参与観察といった複数のデータ収集を実施した。結果、ウォッチャーの語彙テストの達成度がわずかに高く、その理由として、プレイヤーは、英語の学習に比べてタスクに対する精神的労力が大きく、タスクの難易度が高い点が考えられる。一方、アンケートでは、プレイヤーは、英語学習として英語でゲームをプレイすることに積極的であり、また英語学習効果が高いと回答した。しかし、インタビューでは、ゲームスキルの低いプレイヤーはゲームの制御に忙しく英語学習を十分にできず、ウォッチャーは視聴を退屈だと感じたとの回答があった。最後に、今後のフォローアップ研究のための課題を提示した。

Keywords: game-based language learning, watching gameplay, physical interactivity, language learning

Digital game-based language learning (DGBLL) is a research field that explores facilitating language learning through playing computer games, drawing on elements of computer-assisted language learning (CALL) and game-based learning (GBL). DGBLL utilizes inherent game features designed for entertainment that work to increase user motivation. Additionally, certain aspects favorable to facilitating language acquisition are prevalent in games, and playing games has been shown to have several advantages transferrable to learning, such as providing challenge, competition, purpose, and control (Admiraal et al., 2011; Nakamura & Csikszentmihalyi, 2009; Whitton, 2014). Games can also foster learning by means of scaffolding (Sun et al., 2021) and facilitating safe (Jabbari & Eslami, 2019) and immersive environments (Johnsen et al., 2021). Within this controlled environment, feedback is given (Calvo-Ferrer, 2021), and language negotiation can take place through interaction and collaboration (Peterson, 2016). Related studies typically involve users playing games themselves through controllers, but in recent years, watching gameplay has become a popular phenomenon, especially amongst the high school and university age brackets, where almost half (41%) of the popular game streaming website, Twitch.tv's audience is estimated to be between the age of 16 and 24 (Iqbal, 2022). Given the novelty of watching gameplay, its potential for language learning remains underexplored.

Watching gameplay is the act of someone watching another person play a digital game, either live or prerecorded, through websites such as Twitch.tv and YouTube. It is "...a kind of real-time video social media that integrates traditional broadcasting and online gaming" (Li et al., 2020, p. 1). The streamers often provide commentary on the game, and users are able to communicate with the streamer and fellow viewers through a chat box during live sessions. Watching

gameplay has significantly increased in viewership and market value (Hamilton et al., 2014), often rivaling cable TV networks (Gilbert, 2018). And this has only increased during the COVID-19 pandemic (Clement, 2021) as people have sought to alleviate associated negative emotions by turning to video games (Scerbakov et al., 2022). In fact, in some cases, more people watch someone else play a game than play it themselves (Kaytoue et al., 2012; Orme, 2021). This trend is especially taking hold in Japan, as can be seen in a recent survey of middle school students' preferred future careers. For the first time, "Game Streamer" was ranked top five for males and top ten for females (Sony Life Insurance Corporation, 2021). In terms of language education, some streamers dedicate their channels to language learning, such as playing Japanese role-playing games while teaching their audience Japanese. Yet, despite its popularity, we know little about the potential for pedagogical applications in terms of vocabulary achievement, learners' attitudes toward learning a language this way, and their perceptions of its effectiveness.

Watching gameplay shares similarities to media watching, which is not new to education research, as studies have extensively looked at media applications such as movies and TV shows (see Parmawati & Inayah, 2019; Vanderplank, 2019). Nevertheless, traditional media is mostly a passive experience, whereas watching gameplay offers added opportunities for language learning through communicative interactions between viewers and streamers, something that has been shown to be conducive to language learning in game environments (Peterson, 2016).

Literature review

The rise in watching gameplay has encouraged researchers to explain the trend, primarily focusing on its social phenomenon. Typically, studies can be categorized into the subjects of online social interaction (Churchill & Wen Xu, 2016; Diwanji et al., 2020; Hamilton et al., 2014), media consumption (Jang & Byon, 2019; Sjöblom et al., 2017), and motivational appeal (Gros et al., 2017; Sjöblom & Hamari, 2017). Studies on motivational appeal, such as the two listed above, often entail quantitative surveys with predetermined categories offering limited criteria such as entertainment, communicating with others in the online community, checking out a game before purchasing it, and learning game strategies. Notably, 'language learning' is not a selectable option.

Thus, despite the recent studies, there is minimal data on the potential linguistic effectiveness of watching a game versus playing one. And in this regard, the main difference between the two is the utilization or exclusion of game physical interactivity. Interactivity can have many forms, but physical interactivity generally entails a player's physical interaction that they have with the game using a device (deHaan et al., 2010; Sims, 1997). Thus, it can be described as the act of directly manipulating gameplay via a hand-operated controller, something only a player would do while a watcher does not. This is different from social interaction with other human players. When addressing game interactivity, researchers

typically measure its effect on vocabulary acquisition, attention, mental effort and cognitive load, motivation, and flow experience (Cho et al., 2021; deHaan & Kono, 2010; deHaan et al., 2010; Ebrahimzadeh, 2017; Ebrahimzadeh & Alavi, 2016). While studies on game interactivity are not directly related to the watching gameplay phenomenon, they are nonetheless important for determining the relative effectiveness of language learning through watching gameplay.

The evidence differs regarding the effects of game interactivity. When learners are engaged in tasks that elicit a high degree of involvement, language acquisition, such as vocabulary learning, has been known to be facilitated (Peterson, 2021). In support of this notion, Ali Mohsen (2016) conducted a study on vocabulary learning for a serious game designed to teach knee surgery, with results showing the players outperforming the watchers. Nevertheless, interactivity has been negatively associated with causing split attention and overwhelming mental capacity to acquire target learning. This is because operating a physical controller “requires frequent input from the player and the input required can disrupt the player’s involvement with the game space” (Taylor, 2002, p. 20). This is similar to prior studies showing instructional media applications causing split attention and extraneous cognitive load (Kalyuga et al., 1999), factors not desirable for learning. Cognitive load theory states that mental capacity is finite and, when overwhelmed with various mental tasks, can decrease learners’ opportunities for target learning (Sweller, 1994). Following cognitive load theory, physical interactivity would most likely increase mental effort (Pellouchoud et al., 1999) and hamper vocabulary learning (deHaan, 2005). Overall, researchers have stated that the level of cognitive load created by physical input needs to be further addressed (deHaan et al., 2010; Plass & Jones, 2005).

Other studies have inadvertently tested game physical interactivity by assessing the level of technological engagement (LTE) on learning in various education fields outside of language learning. Here, a game-based application is used at the highest level of LTE since it involves users actively engaging with technologies. Secondly, a video-based treatment is implemented at the middle level. Lastly, traditional classroom teaching constitutes no LTE. These studies test whether higher technology engagement leads to better measurable learning results and encompass various subjects such as software engineering (Gordillo et al., 2022; Hsu & Lin, 2016), biology (Chang et al., 2016), science (Chen et al., 2021), intercultural learning (Busse & Krause, 2016), history (Ijaz et al., 2017), and math (Kebritchi et al., 2010; Wang et al., 2018). It is important to note that many of the video-based treatments significantly differ from watching the game-based applications. Instead, participants are usually given a separate form of media with the same target vocabulary.

Distinguishing game types: serious vs. commercial off-the-shelf games

When assessing watching gameplay for language-learning purposes, it is necessary to distinguish the two distinct game types used in DGBLL applications. Commercial off-the-shelf (COTS) games, also known as non-serious games, are

produced primarily for entertainment purposes, and researchers manipulate them to teach designated content in various fields, such as language learning (see deHaan et al., 2010; Ebrahimzadeh, 2017; Peterson, 2016). The other category of games is called serious games, and these are designed primarily for educational purposes. This distinction becomes essential when considering that, while interest in DGBLL research has grown considerably recently, the majority of studies pertain to serious games created by researchers (Boyle et al., 2016; Girard et al., 2013). In accordance with this, the cited studies in the previous paragraph testing game interactivity via LTE can be considered serious games. Conversely, the vast majority of online gameplay watching is done entirely through COTS games.

There are minimal studies looking at game interactivity with COTS games. deHaan (2005) tested Japanese students playing a baseball game, with results indicating students had their attention split between controlling the gameplay and learning the language. The author then conducted two follow-up studies incorporating vocabulary achievement and subjective cognitive load for players and watchers for a reflex-based game (deHaan & Kono, 2010) and a music game (deHaan et al., 2010). Both studies showed a higher vocabulary achievement and less extraneous (negative) cognitive load for the watchers, indicating that watching may have learning advantages over playing games because the watchers can focus on learning without the interruption of a controller. At the time of this current study, the researcher was able to find two more recent experiments, both by the same author utilizing the same game called *Defense of the Ancients*. The studies claim to use a real-time strategy (RTS) game, but this is actually a multi-player online battle arena (MOBA) game made using the RTS game *Warcraft III*. In the first study (Ebrahimzadeh & Alavi, 2016), 136 high school students played or watched the game for five weeks and measured e-learning enjoyment and vocabulary learning. The results showed no significant difference between watchers and players. In the later study (Ebrahimzadeh, 2017), 241 male high school students were put into groups of readers, players, and watchers for five weeks. The findings indicate that players and watchers outperformed the readers. It is important to note that both of these studies included mainly short, text-only dialogues with minimal audio sequences.

In addition to the mentioned studies on COTS games, other researchers have looked at COTS games and the role of interactivity within topics unrelated to education, such as the responsibility and degree of character identification in violent games (Walter & Tsfati, 2016), cognitive load and rape acceptance (Read et al., 2018), violence, perceived difficulty, and frustration (Polman et al., 2008), and motivational processing and cognitive load (Huang, 2011).

Different games may be suitable or inappropriate for language learning (deHaan et al., 2010), so testing various game genres is essential. One important COTS game genre that has not been tested for interactivity is RTS games, a sub-genre of strategy games. Unlike turn-based strategy games, RTS games involve all players moving simultaneously, thus, creating a fast-paced experience with

high player interactivity. RTS games have also been seen to be good for the flow experience (Sweetser & Wyeth, 2005).

Research questions

Given the exploratory nature of the research, this study entails a preliminary study with two purposes. The first purpose is to initially test the effect of physical interactivity on a COTS RTS game. Here, vocabulary achievement and invested mental effort can measure the relative effectiveness between playing and watching a game for language-learning purposes. In addition, this study addresses the potential feasibility for practical applications by soliciting learner attitudes toward playing and watching games for language-learning purposes and their perceptions of its pedagogical effectiveness (see Bolliger et al., 2015). The second purpose is to make observations useful for performing larger studies. The research questions are as follows.

Q1. Does watching gameplay cause higher measurable vocabulary achievement over playing games?

Q2. What are learners' attitudes towards playing or watching games for language-learning purposes, and what are their perceptions about its effectiveness for language learning?

Q3. What effect does the assigned task (play or watch) have on subjective invested mental effort?

Q4. What are learning points that can be considered in a larger experiment?

Method

Experiment design and participants

The experiment was designed to test the effects of game interactivity on vocabulary recall, attitudes and perceptions for pedagogy, and subjective mental effort. Following deHaan et al. (2010), interactivity was controlled by assigning participants into one of two groups: one that would play the game and another that would watch. Interactivity, or the ability to control the gameplay, was assigned as the independent variable. Dependent variables included vocabulary achievement, attitudes and perceptions, and mental effort. To eliminate social interaction influences, participants were not permitted to talk to each other during the treatment. A mixed-methods approach of both quantitative and qualitative data collection was selected based on the recommendations of deHaan et al. (2010) consisting of a vocabulary test, questionnaire, interviews, and researcher observations.

The participants were recruited from one campus at a university in Japan via convenience sampling due to necessity, and advertising was done on the campus electronic bulletin board and via sending emails. In total, 11 participants ($n = 11$) volunteered and took part in the experiment and received payment for their participation (Table 1).

Table 1
Demographics

ID	Gender	Major	Year	English score	Time abroad?	Wkly game play	Wkly game watch
1	F	Engineering	M 1	TOEIC 865	Australia, 2 mo	none	none
2	F	Human Science	M 2	IELTS 6.5	Denmark, 6 mo	none	none
3	M	Engineering	B 1	Beginner	none	none	none
4	M	Engineering	M 1	TOEIC 895	none	none	2-4 hrs
5	M	Human Science	D 2	IELTS 6.0	UK, Canada, 1 yr.	none	none
6	M	Literature	M 1	TOEIC 815	none	< 1 hr	none
7	M	Engineering	M 1	TOEIC 925	Australia, 1 mo.	< 1 hr	5-7 hrs
8	M	Engineering	M 1	TOEIC 770	none	2-4 hrs	2-4 hrs
9	F	Science	B 4	TOEIC 675	UK, 2 wks	none	< 1 hr
10	M	Engineering	B 3	TOEIC 585	none	none	none
11	F	Human Science	B 2	IELTS 6.5	none	none	none

A summary explanation is given as follows. Seven were males, and four were females. Six major in engineering, three in human science, one in science, and one in literature. Six participants are master's students, four are undergraduates, and one is a doctoral student. The participants self-reported their English proficiency scores as follows. Seven gave their TOEIC scores with a mean score of 790. Three elected their IELTS score, including one score of 6.0 and two of 6.5. And one participant wrote "beginner level." Next, the participants were asked about their international experience. Five participants have traveled abroad with a mean time of 17.6 weeks. For game habits, eight don't currently play games, while two play games for less than an hour, and one plays between two and four hours. They mostly play games through game consoles and smartphones. One participant has tried learning English by both playing games and watching gameplay. The typical method of studying English was through university classes, movies and videos, smartphone apps, and textbooks. Next, one participant has had some experience playing RTS games, but none have played the chosen game for this experiment (StarCraft II) before. Five participants have had some experience playing digital games on a PC using a mouse and keyboard.

Resources and procedures

The game selection criteria included a COTS RTS game with a single-player campaign. The following features were prioritized in the selection process, including sufficient audio and text dialogues (related studies typically utilize text-based dialogues only), adequate player physical interactivity during times of dialogue instead of movie-like cinematic cutscenes, and selectable missions for better control. Based on these criteria, *StarCraft II: Wings of Liberty* was chosen. Produced by Blizzard Entertainment and released in 2010, *StarCraft II* has been one of the most popular RTS games, with global multiplayer tournaments currently taking place. This study utilized a mission called *Outbreak*, which was suitable due to its sufficient audio dialogue. Also, as a defense map, it allows a more consistent experiment time control. The mission uses a day/night cycle where players must guard their base during the night and attack the enemy during the day. The mission is won when all surrounding enemy structures are destroyed, and mission failure occurs when all of the players' buildings are destroyed. The game difficulty was set to the lowest level of 'casual,' which is recommended for players with little to no experience playing strategy games.

Two player stations were located adjacent to each other and separated by a divider (Figure 1). The watchers were located out of sight at a nearby station. One laptop and one desktop computer were used, each equipped with an Intel i7 processor and GeForce RTX 3080 graphics card, which exceed the manufacturer's recommended specifications. Both computers were connected to a 27-inch monitor, mouse, keyboard, and stereo headphones. A video and sound splitter were used to run the simultaneous video and audio feed to the watchers. Each watcher used a 27-inch monitor and stereo headphones.

Figure 1

Experiment player station



The participants first completed a demographic questionnaire and inputted their availability. Based on their answers regarding English proficiency level, international experience, and game experience, they were grouped with other participants with similar responses (deHaan et al., 2010). Nevertheless, the researcher needed to also take into consideration the limited sample size and participants' availability. Five groups in total were created, four with two members and one with three members. They were then randomly assigned to the player or watcher role. The experiment needed to be conducted over three days due to the participants' availability and available computers (Figure 2).

Figure 2

Daily schedule for experiment

Day 1	
Participant	Role
1	Player
2	Watcher

Day 2	
Participant	Role
3	Player
4	Watcher

Day 3	
Participant	Role
7	Player
8	Watcher

Day 2	
Participant	Role
5	Player
6	Watcher

Day 3	
Participant	Role
9	Player
10	Watcher
11	Watcher

A week before the experiment, the participants were given prework that they completed at their leisure. This was done to minimize the necessary in-person experiment time and minimize prolonged exposure during COVID-19. The participants completed a vocabulary pre-test and attitudes and perceptions questionnaire and watched a 15-minute video on YouTube in Japanese explaining the core mechanics of the game and how to play the Terran (human) race, which the game's single-player campaign focuses on.

Efforts were made to minimize the in-person treatment time to avoid participants becoming bored (deHaan et al. 2010). Nevertheless, it was also necessary to ensure that the players were given a sufficient understanding of the complex mechanics of the game. The target experiment time was set to under 90 minutes, including the game tutorial (10 mins), 1v1 skirmish map (10 mins), mission (40 mins), post-test and questionnaire (20 mins), and interviews (10 mins). When participants arrived, they were given an explanation sheet in Japanese outlining the experiment scope, data collection, and other details. They were given the opportunity to ask questions before giving their informed consent.

The participants were then given a schedule sheet explaining each step of the experiment and assigned to their stations. The instructions were to either play or watch the mission and try to learn the language. No interaction with other participants, notetaking, or word searching were permitted. The players were not permitted to pause the game unless necessary. And in the event of mission failure or success, they were asked to await further instructions from the researcher regarding repeating the mission or finishing the experiment.

The session began with a tutorial scenario with English explanations of the game mechanics since the game does not have a Japanese language option. The participants were informed that they did not need to learn the English for the tutorial and subsequent practice skirmish scenario. Instead, these steps were only there to allow them to learn about the game. After the tutorial, the players were

allowed to do a one vs. one skirmish against a computer opponent (Terran vs. Terran) for 10 minutes on the lowest opponent difficulty (very easy) to try the game out for themselves. Once again, the watchers watched the player play this section. The participants were then informed that the experiment mission would begin, where they would play or watch the game while trying to learn the English.

To ensure a smooth and consistent treatment, the researcher acted as an active participant when necessary in certain cases, including when participants asked questions or for scenario situations (mission success or failure). Internal rules were established that the participants were not aware of, explained as follows. For mission failure, a threshold time of 45 minutes was established. If the player failed the mission prior to this time, they would be instructed to restart the mission. If this time had passed, the player would restart the mission with cheats enabled, making the game easier by increasing the players' units' life points and attack damage. If the player failed with this mode, the experiment would then be over. In accordance with this, three groups had cheats enabled, and they were then able to complete the mission.

Data collection and analysis

Vocabulary test

A written vocabulary test was created for this study and administered twice, prior to the treatment and immediately after. The creation of the test is explained as follows. Multiple-choice questions were avoided to prevent the participants from deducing the correct answer based on the given choices or simply guessing the correct answer. Instead, the participants were given an entire sample sentence in English taken from the mission script with the target word or phrase in bold. They were asked to translate only the bold section into Japanese by writing it in a text box. If they did not know the answer, they could write, 'I don't know.' A correct/incorrect system was used by assigning 1 point for correct answers and 0 for incorrect answers. The results were scored individually by the author and a native Japanese speaker with advanced English proficiency. Any discrepancies were then discussed by the two graders, and a score was mutually agreed upon.

The following steps were taken for selecting the test items. First, the entire English dialogue of the selected mission was first written out. Then individual words were selected as well as compound words such as phrasal verbs and compound nouns, excluding colloquial words or idioms that require cultural context to understand. The selection of the items was based on the following criteria. First, words relevant to the mission scenario and means of accomplishing the mission were prioritized. Additionally, there was a focus on choosing dialogue that is observable in that the prompt was given, followed by the player or watcher immediately seeing its action played out on the screen, allowing them to deduce the meaning even if they don't know the word. Finally, low-frequency words that were most likely unknown to the participants based on the author's assessment were selected. The presentation order and frequency were controlled by the game and the player. No control could be implemented on the number of times

each word was witnessed. Nevertheless, grouping the players ensured a consistent experience between each player and their associated watcher. Altogether, 33 target items were chosen, and in order to prevent 'prepping' the participants, ten unrelated distracter words were created, along with their associated sentences. Thus, the total question count was 43, arranged in random order. The participants were instructed to finish the vocabulary test in 10 minutes and not to use dictionaries or internet searches to find the answers.

Attitudes and perceptions questionnaire

The same attitudes and perceptions questionnaire was conducted before and immediately after the treatment. It consisted of 25 Likert items derived from prior surveys, as existing questionnaires can be considered more reliable as they have been tested for validity and reliability (Alqurashi, 2016). The survey solicited attitudes and perceptions toward playing games and watching gameplay for language learning purposes. This was derived from Bolliger et al. (2015) ($\alpha = .72$) and Bourgonjon et al. (2010) ($\alpha > .70$), who tested students' perceptions and attitudes regarding playing digital games. These questionnaires were based on the traditional Technology Acceptance Model (TAM) by Davis (1989) and include the four criteria of ease of use, learning opportunities, experience, and preference. As with Bolliger et al. (2015), the survey included a 4-point scale ranging from 1 (strongly disagree) to 4 (strongly agree) to prevent the participants from choosing a neutral option, as has been an issue with Japanese students (Wang et al., 2008). The items were translated from English into Japanese by a native Japanese speaker with advanced proficiency in English.

Invested mental effort and perceived difficulty

The participants answered four questions after the treatment regarding their subjective invested mental effort and perceived difficulty of their assigned task. These questions were based on deHaan et al. (2010) (mental effort $\alpha = .551$, material difficulty $\alpha = .565$), who used the prior surveys of Kalyuga et al. (1998) ($\alpha = .4583$) and Paas (1992) ($\alpha > .85$), and are shown as follows.

Q1. How much mental effort did you invest in your assigned task (playing/watching)? [9-point Likert scale ranging from extremely low to extremely high mental effort]

Q2. How much mental effort did you put into learning English through your assigned task (playing/watching)? [9-point Likert scale ranging from extremely low to extremely high mental effort]

Q3. How difficult was your assigned task (playing/watching)? [7-point Likert scale ranging from extremely easy to extremely difficult]

Q4. How difficult was it to understand the English in the game? [7-point Likert scale ranging from extremely easy to extremely difficult]

Interviews

The post-treatment interviews were conducted individually in Japanese or English and structured on the questionnaire items (attitudes and perceptions, invested mental effort, perceived difficulty) (see Appendix). Nevertheless, they allowed the flexibility for participants to elaborate on their overall ideas. The conversations were recorded with permission and later transcribed and translated into English by a native Japanese speaker with advanced English proficiency. The transcripts were then coded to reveal trends.

Researcher observations

During the experiment, the researcher took notes on the physical actions of the participants (e.g., body posture, eye attentiveness, and signs of fatigue) and the in-game decisions the players were making throughout the mission. These were collected to offer suggestions for future studies regarding running them smoothly.

Results

Quantitative data

The vocabulary test and invested mental effort questions assess comparative effectiveness while the perceptions and attitudes survey examines participants' feelings towards pedagogy.

Vocabulary Test

The descriptive statistics of the pre and post-vocabulary tests are displayed in Tables 2 and 3.

Table 2

Descriptive statistics of the vocabulary pre-test

	N	Min	Max	Mean	Std. deviation
Players	5	9	20	13.8	4.02
Watchers	6	13	21	15.17	3.97

*max possible score = 33

Table 3

Descriptive statistics of the vocabulary post-test

	N	Min	Max	Mean	Std. deviation
Players	5	13	19	16.4	2.19
Watchers	6	14	26	19.67	3.89

*max possible score = 33

Both groups scored similarly in the pre-test, with the watchers having a slightly higher mean score (watchers = 15.17, players = 13.8). The results for the post-test show both groups increased in means. The watchers' increase in points was slightly higher (mean increase of 4.5) than the players (mean increase of 2.6). Additionally, while the watchers group increased both their minimum and maximum points, the players group had lower minimum and maximum scores on the post-test. In these cases, the participants gave correct answers on the pre-test but incorrectly changed them after the treatment.

Attitudes and perceptions

The results of the attitudes and perceptions questions are displayed in Tables 4 and 5. The players' mean scores increased slightly in all three categories (ease of use, learning opportunities, preference) after the treatment. Conversely, the watchers had the same mean score for ease of use, and both learning opportunities and preference decreased post-treatment.

Table 4

Means and standard deviations for perceptions and attitudes (players)

Item		M	SD
Ease of use	before	2.10	0.65
	after	2.30	0.27
Learning opportunities	before	2.65	0.38
	after	2.85	0.29
Preference for learning English this way	before	2.00	0.41
	after	2.33	0.47

*n = 5; max is 4

Table 5

Means and standard deviations for perceptions and attitudes (watchers)

Item		M	SD
Ease of use	before	2.42	1.07
	after	2.42	0.66
Learning opportunities	before	2.30	0.72
	after	2.40	0.96
Preference for learning English this way	before	2.27	1.00
	after	2.17	0.92

*n = 6; max is 4

Invested mental effort and difficulty

The results of the participants' subjective invested mental effort and difficulty of the task are displayed in Table 6. The players experienced a higher mental effort for their assigned task of playing (M = 7.00, SD = 1.58) than the watchers (M = 5.17, SD = 2.04). They also had less mental effort dedicated toward learning the English (M = 4.00, SD = 1.58) than the watchers (M = 6.00, SD = 2.37). Finally, the players felt their assigned task was more difficult (M = 5.60, SD = 0.55) than the watchers (M = 4.67, SD = 1.37), and the players felt they had slightly more difficulty understanding the English (M = 5.4, SD = 0.89) compared to the watchers (M = 5.17, SD = 1.17).

Table 6

Means and standard deviations for invested mental energy

Item		M	SD
Q1 Mental effort of assigned task	player	7.00	1.58
	watcher	5.17	2.04
Q2 Mental effort in learning English	player	4.00	2.24
	watcher	6.00	2.37
Q3 Difficulty of assigned task	player	5.60	0.55
	watcher	4.67	1.37
Q4 Perceived difficulty of understanding the English	player	5.4	0.89
	watcher	5.17	1.17

*Q1 & Q2: max = 9

*Q3 & Q4: max = 7

Qualitative data

The interviews and researcher observations provided additional explanation to the above quantitative data.

Interviews

The eleven interviews lasted approximately 15 minutes and yielded the following trends.

Trend 1: Most players and watchers feel watching is more effective for learning English.

Four of the five players felt watching is more effective for learning English. One stated, "For sure I could have learned more English if I was watching today's game." Another reiterated, "When it comes to playing the game, it was quite difficult to learn the English since I had to concentrate on playing. Overall, ... I think it's more effective to watch it." The one player who felt playing is more effective explained, "if I watch the games, maybe I cannot focus on the games. Maybe I will ... (get distracted)." Nevertheless, all five reported difficulties concentrating on the language because they were too busy playing. One participant elaborated, "I couldn't understand what it was saying, and it was difficult for me to understand it properly because I was distracted by the game." Even the player who felt playing is more effective said, "I didn't care so much about the English. Just attack." One player made an interesting observation when he likened this experience to a flight school training program he is currently undergoing. One task involves him taking turns between flying a plane and watching a fellow trainee fly it in a computer simulation. The directions are simultaneously explained in English, and he described the experience by saying, "... when I watch other students do the simulation from behind, the content of the flowing English is absorbed quickly ... but when I was actually holding the control stick, I was desperate to control it, and no (English) was absorbed in the end."

For the watchers, two of the six felt watching is more effective for learning English. One stated, "I think it's better to just watch while studying English. Overall, if you play the game yourself, you have to think about how to play the game. But if you just watch it, you will watch the game with the intention of listening to English and understanding it." Another agreed, "Well, I can concentrate on the subtitles in English. If I do the game, I think I can't focus on the English subtitles." Additionally, one watcher felt both watching and playing are effective for people interested in games, while another had mixed feelings. Only one watcher felt playing is more effective.

Trend 2: The skill of the player determines the viewing experience of the watcher.

Both the watcher who felt playing is better for learning English and the one with mixed feelings reported their experience being ruined by boredom from the player having low skill. "It's a little bit irritating because this was a strategy game but the player ... had no strategy," one complained. "I think I know better than him regarding the rules." The other stated, "I was more tired than expected ... I felt like

it was repeating the same thing, and I was a little sleepy.” In these cases, the mission runtime ran over an hour.

In addition to the trends, other notable findings recorded from the interviews are listed as follows. Two participants regularly watch gameplay but don't play games themselves, an occurrence found in other studies, as mentioned previously (Kaytoue et al., 2012). One player felt that it would have been more effective if he could pause the game to look up unknown words.

Researcher observations

The following observations were found by the researcher regarding areas of potential improvement for future follow-up studies.

Observation 1: Players require more training for the game beforehand. Despite receiving various opportunities to learn the game (online tutorial video, in-game tutorial scenario, and 1v1 skirmish scenario with a computer player), some players showed signs that they could not fully operate the game. Some observed examples of the players' actions include: not realizing the game's 'fog-of-war' mechanic, not creating Supply Depots to increase the maximum army size, relying on mouse clicks only without using keyboard shortcuts, and spending significant amounts of time before attacking the enemy base.

Observation 2: Participants appeared bored when the experiment runtime was long. In addition to some watchers complaining in the interviews of mental exhaustion and boredom from players' low skill level, the researcher also observed participants from both groups stretching their arms and backs and yawning multiple times, especially when the mission ran for over an hour. Selecting this defense-type mission did not ensure a consistent experience as intended since it would not end until the enemy base was destroyed. The researcher underestimated the time needed for players to achieve the objective, and consequently, the treatment runtime exceeded the intended time.

Discussion

The results of the vocabulary test show that the watchers performed slightly better than the players in the mean total score and the minimum and maximum scores. This supports the findings found in reflex games (deHaan & Kona, 2010) and music games (deHaan et al., 2010) while differing from the serious game used by Ali Mohsen (2016). And, since both groups performed better on the post-test, the results may also provide some support for Ebrahimzadeh (2017) despite this current study not implementing a reading group for comparison.

Next, the results of subjective invested mental effort also support the findings of previous studies (deHaan & Kona, 2010; deHaan et al., 2010) in that the players reported comparatively higher mental effort put into their tasks, less mental effort allocated to learning the English, and rated their task as more difficult. When

combining this result with the higher vocabulary scores for the watchers, this study supports assertions that physical interactivity may have an adverse effect on vocabulary learning (deHaan et al., 2010) in RTS games.

Furthermore, the attitudes and perceptions survey shows that the players' feelings improved after the treatment while the watchers' scores stayed the same or decreased. Yet, the interviews contradicted this in revealing that most players felt watching is more effective for learning English, while the watchers were less certain. One explanation for this may be that, as stated by two of the watchers, the lack of skill or insufficient preparation time of the player may have negatively influenced the watchers' overall experience. This was also seen in the researchers' observations, as some players appeared to not have adequately understood the game mechanics despite receiving prior preparation time and playing the game at the lowest difficulty setting. Overall, the players felt overwhelmed while some of the watchers felt bored, and this may have led to different results had the players been allowed more time to master the complex mechanics apparent in RTS games. Indeed, the viewership for online game watching is significantly higher for professionals than amateurs, and many viewers watch tournaments to witness the best players compete. This has not been considered in related studies testing interactivity and may have implications for studies and merit future investigation. Experiments also may need to be adjusted based on their affiliated game's learning curve to ensure an appropriate level of functionality that does not interfere with the experiment results.

Conclusion

Watching gameplay has become an international sensation, which opens up the importance of considering its merits for pedagogy and the effects of physical interactivity on COTS games. This study was conducted as an exploratory and preliminary step for testing the physical interactivity of an RTS game. The findings indicate adverse effects on interactivity, raising the prospect of the potential for watching gameplay for language-learning purposes.

While this study provides initial evidence for an RTS game, there are several limitations that prevent us from making claims or generalizations. Mainly, the sample size is small, and the treatment and test were conducted only one time. Additionally, different game genres have varying degrees of player interactivity and thus may yield different results. Therefore, future experiments can consist of larger sample sizes and longitudinal studies with delayed post-tests (Peterson, 2010) testing different genres (deHaan et al., 2010). It is hoped that the results of this study can assist in developing such studies.

References

- Admiraal, W., Huizenga, J., Akkerman, S., & Ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior*, 27(3), 1185–1194. <https://doi.org/10.1016/j.chb.2010.12.013>
- Ali Mohsen, M. (2016). The use of computer-based simulation to aid comprehension and incidental vocabulary learning. *Journal of Educational Computing Research*, 54(6), 863–884. <https://doi.org/10.1177/0735633116639954>
- Alqurashi, M. A. (2016). Saudi teachers' experiences and attitudes toward integrating video games for learning: Affordances and constraints of using video games in Saudi Arabian classrooms [Doctoral dissertation, University of Northern Colorado] *UNC Dissertations database*. <https://digscholarship.unco.edu/cgi/viewcontent.cgi?article=1360&context=dissertations>
- Bolliger, D. U., Mills, D., White, J., & Kohyama, M. (2015). Japanese students' perceptions of digital game use for English-language learning in higher education. *Journal of Educational Computing Research*, 53(3), 384–408. <https://doi.org/10.1177/0735633115600806>
- Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Computers & Education*, 54(4), 1145–1156. <https://doi.org/10.1016/j.compedu.2009.10.022>
- Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, 94, 178–192. <https://doi.org/10.1016/j.compedu.2015.11.003>
- Busse, V., & Krause, U. M. (2016). Instructional methods and languages in class: A comparison of two teaching approaches and two teaching languages in the field of intercultural learning. *Learning and Instruction*, 42, 83–94. <https://doi.org/10.1016/j.learninstruc.2016.01.006>
- Calvo-Ferrer, J. R. (2021). Effectiveness of Type of Feedback and Frequency on Digital Game-Based L2 Vocabulary Acquisition. *International Journal of Game-Based Learning (IJGBL)*, 11(3), 38–55. https://rua.ua.es/dspace/bitstream/10045/116614/1/Calvo-Ferrer_2021_IntJGame-BasedLearn_final.pdf
- Chang, C., Chang, C. K., & Shih, J. L. (2016). Motivational strategies in a mobile inquiry-based language learning setting. *System*, 59, 100–115. <https://doi.org/10.1016/j.system.2016.04.013>

- Chen, Y. C., Lu, Y. L., & Lien, C. J. (2021). Learning environments with different levels of technological engagement: a comparison of game-based, video-based, and traditional instruction on students' learning. *Interactive Learning Environments*, 29(8), 1363–1379.
<https://doi.org/10.1080/10494820.2019.1628781>
- Cho, J. Y., Andersen, E., & Kizilcec, R. F. (2021). Delivery Ghost: Effects of Language Immersion and Interactivity in a Language Learning Game. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–7. https://dl.acm.org/doi/pdf/10.1145/3411763.3451767?casa_token=9q0bSxZ65MwAAAAA:FHmX0tMSeaBuxHQ82gh-QysAqdl5UI2rMsGcDcSI2wtBSsWoKJFrDSDjOfgSOYUcj9mo2xX3epE8RA
- Churchill, B. C., & Xu, W. (2016). The modern nation: A first study on Twitch. tv social structure and player/game relationships. In 2016 *IEEE international conferences on big data and cloud computing (BDCloud), social computing and networking (SocialCom), sustainable computing and communications (SustainCom)(BDCloud-SocialCom-SustainCom)*, 223–228.
<https://doi.org/10.1109/BDCloud-SocialCom-SustainCom.2016.43>
- Clement, J. (2022, November 16). *Average number of concurrent viewers on Twitch from 2nd quarter 2018 to 3rd quarter 2022*. Statista. <https://www.statista.com/statistics/761122/average-number-viewers-on-youtube-gaming-live-and-twitch/>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
<http://dx.doi.org/10.2307/249008>
- deHaan, J. (2005). Acquisition of Japanese as a foreign language through a baseball video game. *Foreign Language Annals*, 38(2), 278–282.
<https://doi.org/10.1111/j.1944-9720.2005.tb02492.x>
- deHaan, J., & Kono, F. (2010). The effect of interactivity with WarioWare Minigames on second language vocabulary learning. *Journal of Digital Games Research*, 4(2), 47–59.
- deHaan, J., Reed, W. M., & Kuwanda, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, 14(2), 74–94. <https://www.lltjournal.org/item/543/>
- Diwanji, V., Reed, A., Ferchaud, A., Seibert, J., Weinbrecht, V., & Sellers, N. (2020). Don't just watch, join in: Exploring information behavior and copresence on Twitch. *Computers in Human Behavior*, 105, 1–11.
<https://doi.org/10.1016/j.chb.2019.106221>
- Ebrahimzadeh, M. (2017). Readers, players, and watchers: EFL students' vocabulary acquisition through digital video games. *English Language Teaching*, 10(2), 1–18. <https://eric.ed.gov/?id=EJ1125224>

- Ebrahimzadeh, M., & Alavi, S. (2016). Motivating EFL students: E-learning enjoyment as a predictor of vocabulary learning through digital video games. *Cogent Education*, 3(1), 1–14.
<https://doi.org/10.1080/2331186X.2016.1255400>
- Gilbert, B. (2018, February 14). *Amazon's streaming service Twitch is pulling in as many viewers as CNN and MSNBC*. Business insider.
<https://www.businessinsider.com/twitch-is-bigger-than-cnn-msnbc-2018-2>
- Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? a meta-analysis of recent studies. *Journal of Computer Assisted Learning*, 29(3), 207–219.
<https://doi.org/10.1111/j.1365-2729.2012.00489.x>
- Gordillo, A., López-Fernández, D., & Tovar, E. (2022). Comparing the effectiveness of video-based learning and game-based learning using teacher-authored video games for online software engineering education. *IEEE Transactions on Education*, 65(4), 524–532.
<https://doi.org/10.1109/TE.2022.3142688>
- Gros D., Wanner B., Hackenholt A., Zawadzki P., & Knautz K. (2017). World of streaming. Motivation and gratification on Twitch. In G. Meiselwitz (Ed.), *Proceedings of SCSM 2017: Social computing and social media. Human behavior*. (pp. 44–57). Springer.
https://doi.org/10.1007/978-3-319-58559-8_5
- Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on Twitch: Fostering participatory communities of play within live mixed media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1315–1324). Association of Computing Machinery.
<https://doi.org/10.1145/2556288.2557048>
- Hsu, W. C., & Lin, H. C. K. (2016). Impact of applying WebGL technology to develop a web digital game-based learning system for computer programming course in flipped classroom. *Proceedings of the 2016 International Conference on Educational Innovation through Technology (EITT)* (pp. 64–69). IEEE. <https://doi.org/10.1109/EITT.2016.20>
- Ijaz, K., Bogdanovych, A., & Trescak, T. (2017). Virtual worlds vs books and videos in history education. *Interactive Learning Environments*, 25(7), 904–929. <https://doi.org/10.1080/10494820.2016.1225099>
- Iqbal, Mansoor (2022, September 6). *Twitch Revenue and Usage Statistics (2022)*. Business of Apps. [https://www.businessofapps.com/data/twitch-statistics/#:~:text=to%202021%20\(bn\)-,Twitch%20age%20demographics,are%20aged%2055%20or%20above](https://www.businessofapps.com/data/twitch-statistics/#:~:text=to%202021%20(bn)-,Twitch%20age%20demographics,are%20aged%2055%20or%20above)
- Jabbari, N., & Eslami, Z. R. (2019). Second language learning in the context of massively multiplayer online games: A scoping review. *ReCALL*, 31(1), 92–113. <https://doi.org/10.1017/S0958344018000058>
- Jang, W. W., & Byon, K. K. (2019). Antecedents and consequence associated with esports gameplay. *International Journal of Sports Marketing and Sponsorship*, 21(1), 1–22. <https://doi.org/10.1108/IJSMS-01-2019-0013>

- Johnsen, H. M., Briseid, H. S., Brodtkorb, K., Slettebø, Å., & Fossum, M. (2021). Nursing students' perceptions of combining hands-on simulation with simulated patients and a serious game in preparing for clinical placement in home healthcare: a qualitative study. *Nurse Education Today*, 97, 104675. <https://doi.org/10.1016/j.nedt.2020.104675>
- Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human factors*, 40(1), 1–17. <https://doi.org/10.1518/001872098779480587>
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13(4), 351–372. [https://doi.org/10.1002/\(SICI\)1099-0720\(199908\)13:4<351::AID-ACP589>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6)
- Kaytoue, M., Silva, A., Cerf, L., Meira, Jr., W., & Raïssi, C. (2012). Watch me playing, I am a professional: a first study on video game live streaming. *Proceedings of the 21st International Conference on World Wide Web* (pp. 1181–1188). Association for Computing Machinery. <https://doi.org/10.1145/2187980.2188259>
- Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, 55(2), 427–443. <https://doi.org/10.1016/j.compedu.2010.02.007>
- Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and Higher Education*, 8(1), 13–24. <https://doi.org/10.1016/j.iheduc.2004.12.001>
- Li, Y., Wang, C., & Liu, J. (2020). A systematic review of literature on user behavior in video game live streaming. *International Journal of Environmental Research and Public Health*, 17(9), 1168–1177. <https://doi.org/10.3390/ijerph17093328>
- Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In C. R. Snyder & S. J. Lopez (Eds.), *Oxford handbook of positive psychology* (2nd ed., 195–206). Oxford University Press.
- Orme, S. (2021). “Just watching”: A qualitative analysis of non-players' motivations for video game spectatorship. *New Media & Society*, 24(10), 2252–2269. <https://doi.org/10.1177/1461444821989350>
- Parmawati, A., & Inayah, R. (2019). Improving students' speaking skill through English movie in scope of speaking for general communication. *Eltin Journal: Journal of English Language Teaching in Indonesia*, 7(2), 43–53.
- Pellouchoud, E., Smith, M. E., McEvoy, L., & Gevins, A. (1999). Mental effort related EEG modulation during video game play: Comparison between juvenile epileptic and normal control subjects. *Epilepsia*, 40(4), 38–43. <https://doi.org/10.1111/j.1528-1157.1999.tb00905.x>
- Peterson, M. (2010). Computerized games and simulations in computer-assisted language learning: A meta-analysis of research. *Simulation & Gaming*, 41(1), 72–93. <https://doi.org/10.1177/1046878109355684>

- 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.
<https://doi.org/10.1080/09588221.2016.1197949>
- Peterson, M. (2021). Digital simulation games in CALL: a research review. *Computer Assisted Language Learning*, 1–24.
<https://doi.org/10.1080/09588221.2021.1954954>
- Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *Cambridge handbook on multimedia learning* (pp. 467–488). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511816819.030>
- Polman, H., De Castro, B. O., & van Aken, M. A. (2008). Experimental study of the differential effects of playing versus watching violent video games on children's aggressive behavior. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 34(3), 256–264.
<https://doi.org/10.1002/ab.20245>
- Read, G. L., Lynch, T., & Matthews, N. L. (2018). Increased cognitive load during video game play reduces rape myth acceptance and hostile sexism after exposure to sexualized female avatars. *Sex Roles*, 79(11), 683–698.
<https://doi.org/10.1007/s11199-018-0905-9>
- Scerbakov, A., Pirker, J., & Kappe, F. (2022). When a pandemic enters the game: The initial and prolonged impact of the COVID-19 pandemic on live-stream broadcasters on Twitch. *Proceedings of the Hawaii International Conference on System Sciences HICSS 2022* (pp. 1–10). University of Hawaii.
http://jpirker.com/wp-content/uploads/2022/01/paper_final_manuscript.pdf
- Sims, R. (1997). Interactivity: A forgotten art? *Computers in Human Behavior*, 13, 157–180. [https://doi.org/10.1016/S0747-5632\(97\)00004-6](https://doi.org/10.1016/S0747-5632(97)00004-6)
- Sjöblom, M., Törhönen, M., Hamari, J., & Macey, J. (2017). Content structure is king: An empirical study on gratifications, game genres and content type on Twitch. *Computers in Human Behavior*, 73, 161–171.
<https://doi.org/10.1016/j.chb.2017.03.036>
- Sony Life Insurance Corporation. (2021, July 29). *Chūkōsei ga omoiegaku shōrai ni tsuite no ishiki chōsa [Awareness survey about the future envisioned by junior and senior high school students]*. Sony Life Insurance Corporation.
https://www.sonylife.co.jp/company/news/2021/nr_210729.html
- Sun, L., Ruokamo, H., Siklander, P., Li, B., & Devlin, K. (2021). Primary school students' perceptions of scaffolding in digital game-based learning in mathematics. *Learning, Culture and Social Interaction*, 28, 100457.
<https://doi.org/10.1016/j.lcsi.2020.100457>
- Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, 3(3), 1–24.
<https://doi.org/10.1145/1077246.1077253>

- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295–312.
[https://doi.org/10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5)
- Taylor, L. (2002). *Video games: Perspective, point-of-view, and immersion* [Master's thesis, University of Florida, USA]. Semantic Scholar.
<http://purl.fcla.edu/fcla/etd/UFE1000166>
- Vanderplank, R. (2019). 'Gist watching can only take you so far': Attitudes, strategies and changes in behaviour in watching films with captions. *The Language Learning Journal*, 47(4), 407–423.
<https://doi.org/10.1080/09571736.2019.1610033>
- Walter, N., & Tsfaty, Y. (2016). Interactive experience and identification as predictors of attributing responsibility in video games. *Journal of Media Psychology*, 30(1), 1–32. <https://doi.org/10.1027/1864-1105/a000168>
- Wang, S. Y., Chang, S. C., Hwang, G. J., & Chen, P. Y. (2018). A microworld-based role-playing game development approach to engaging students in interactive, enjoyable, and effective mathematics learning. *Interactive Learning Environments*, 26(3), 411–423.
<https://doi.org/10.1080/10494820.2017.1337038>
- Wang, R., Hempton, B., Dugan, J. P., & Komives, S. R. (2008). Cultural differences: Why do Asians avoid extreme responses? *Survey Practice*, 1(3), 1–8. <https://doi.org/10.29115/SP-2008-0011>
- Whitton, N. (2014). *Digital games and learning: Research and theory* (1st ed.). Routledge.

Appendix

Interview questions framework

Note: This template was used as a guide to ensure the dialogue stayed on topic. Nevertheless, the interviewees were allowed to elaborate on their ideas, and the researcher could pursue other lines of inquiry.

Overall experience and preference

- How was the overall experience today? Do you feel your English level has improved?
- If given a choice, which task (play/watch) would you have liked to have been assigned today?
- Do you play video games or watch gameplay?
- What is your usual way of studying English?

Ease of use (accessibility)

- If you were to continue doing this method (play/watch) by yourself to learn English, how difficult would it be?

Learning opportunities (perceived effectiveness)

- Before the experiment, how effective did you think your assigned task (playing/watching) was for learning English? Has this changed?
- Which (play/watch) is more effective for learning English?

Motivation and attitudes

- Before the experiment, how motivated were you to try your assigned method (playing/watching) to learn English? Did your opinion change?
- How motivated are you to continue learning English this way?
- Do you plan on continuing to study English this way?

Perceived difficulty and mental effort

- How difficult did the game itself seem (unrelated to language learning)?
- How difficult was your assigned task (play or watch while learning English)?
- How much of your mental concentration was used on performing your task versus learning English?
- Which task (play/watch) do you think would generally demand more of your mental effort?

Author's bio

Shawn Andersson is an English Lecturer within the Language Education Center at the Biwako Kusatsu Campus of Ritsumeikan University. His research interests include Game-based Language Learning, Computer-assisted Language Learning, and Self-access Language Learning.

2

Using online forums to promote Japanese university student motivation to use English outside the classroom

Jack Pudelek, Kwansai Gakuin University

Abstract

The aim of this study was to investigate whether online forums, used as a tool for virtual exchange of communication between English learners in different institutions, areas and countries, are beneficial in increasing Japanese University students' motivation to use English outside the classroom. Online forums or message boards have long been used as a means for people around the world to discuss various topics and interact with each other about subjects they share an interest in. These forums can create lively and varied discussions as users comment on and post questions about other users' posts. A forum called the International Virtual Exchange (IVE) Project allows students learning English in universities around the world to take part in an 8-week online English exchange program, discussing four topics by making posts connected to the topics. The topics are related to sharing information and learning about culture and cultural differences. This paper explores whether this activity increases motivation to communicate in the target language outside of the classroom, either instrumentally through the students striving to complete the immediate practical goals of making their posts and interacting with their peers or integratively through providing opportunities for personal growth and cultural enrichment by learning about different cultures through the medium of English (Gardner & Lambert, 1972) or by nurturing an identity connected to the second language.

本研究では地域・国の英語学習者間の仮想的なコミュニケーション交流のツールとして利用されているオンラインフォーラムが日本の大学生の教室外での英語使用に対するモチベーションを高めるために有益かを調査しました。オンライン・フォーラムや掲示板は、世界中の人々が様々なトピックについて議論し、互いに興味を持つテーマについて交流する手段として、長い間利用されてきました。このようなフォーラムでは、ユーザーが他のユーザーの投稿にコメントしたり、質問を投稿したりすることで、活発で多彩な

議論が展開されます。IVE (International Virtual Exchange) プロジェクトと呼ばれるフォーラムでは、世界中の大学で英語を学ぶ学生が8週間のオンライン英語交流プログラムに参加し、4つのトピックに関連した投稿を行いながらディスカッションを行うことができるようになっています。テーマは、情報の共有や文化、文化の違いを学ぶことに関連するものです。本稿では、この活動が教室外でのターゲット言語のコミュニケーションに対するモチベーションを高めるかどうかを調査している。目先の実用的な目標をクリアしようと努力する学生の道具的な動機付けでしょうか。それとも、英語を媒介として異文化を学び、自己成長や文化的な豊かさを得る機会を提供することで、統合的な動機付け(Gardner & Lambert, 1972)を行うのでしょうか。それとも、第二言語と結びついたアイデンティティを育ませるのでしょうか。

Keywords: Virtual exchange (VE), computer mediated communication (CMC), motivation, learner identity

Introduction

Learning a language is a long and arduous endeavor that takes place over numerous years (Klein, 1991). N. Ellis (2002) stated that frequency is a determining factor of language acquisition because the rules of language are structured by the learner through analysis of input over a long period of time. It is widely accepted that second language learners require plentiful opportunities to produce varied output. Swain's (1985) output hypothesis explained the three functions of output are: to notice gaps in knowledge, test tacit hypothesis of previously learnt language and finally metalinguistic reflection, which allows us to internalize linguistic knowledge. Stereotypical foreign language learners in classrooms are exposed to the second language for only a few hours per week (Lightbown & Spada, 2013, p. 38). If you consider the number of hours of exposure to input or opportunities to test output first language learners receive in a natural setting, the few hours per week in the classroom seem rather insufficient. Therefore, it is incumbent that language learners must strive to put in the hours of work necessary to learn a language, seeking out opportunities to use it wherever possible, both inside and outside of the classroom and use these opportunities effectively.

Recent research has shown that adult English foreign language learners' motivation to use English outside of the classroom increases when they are given access to online tools (Aydin, 2007; Columbo & Columbo, 2007; Son, 2007, as cited in Terrell, 2013). One online tool that can help improve saliency in recognizing target language forms is computer mediated communication (CMC) (Smith, 2004). CMC can also increase opportunities to receive feedback and produce more modified output than conventional classroom communication with a teacher (Mackey et al., 2003). A Virtual Exchange (VE) describes a program where CMC is implemented to promote interaction between learners around the world (O'Dowd, 2018, as cited in Canals, 2020). Realtime text, oral or video-based CMC can be referred to as synchronous computer mediated communication (SCMC). Darhower (2008) found that a telecollaborative chat (SCMC) exchange between 80 Spanish and English learners resulted in improvements in their L2 writing and knowledge of the L2 culture. In addition, the learners said that they intended to continue

to communicate with native speakers (NS) using the L2 in the future. Jung et al. (2019) used video SCMC interactions between English learners in Korea, Japan and Taiwan to nurture intercultural and linguistic awareness. There was a positive response from learners who reported that the VE had fostered their L2 learning, intercultural communicative competence and willingness to communicate and had also instilled in them a need to use English.

However, more research into exactly how these virtual exchanges motivate students is needed. What types of CMC work best? Is the motivation instrumental in terms of striving to complete a short term goal or more intrinsic in terms of a purpose of long-term betterment?

This study is an attempt to address these lacunae in the research. It focuses on the use of a text based online forum for Virtual Exchange (VE) between non-native speakers learning English. It attempts to discover if it is effective as a means of motivating out-of-class English practice and if so how. This study was specifically designed to answer the following questions:

RQ1. Can virtual exchanges using text-based communication in online forums such as the International Virtual Exchange Project motivate students to use English outside of the classroom?

RQ2. If so, is the motivation instrumental because of a short-term goal such as completing a homework task and passing the course or is it intrinsic, meaning they are motivated by the idea of being able to communicate in the L2 with people from other countries and learn about their culture? Or is the motivation related to the building of their L2 identity, learner autonomy or other factors?

Literature review

Acquisition of a language requires a focused use of the language over time both inside and outside of the classroom (Hyland, 2004; Nunan & Richards, 2015). The learner seeking the chance to learn out-of-the classroom is one of four factors that Nunan & Richards (2015) attribute to language learning success, along with motivation, goal setting and reflection on progression. Ellis (2008, p. 855) also stated that, in order for language learners to be successful they must not only actively attempt to recognize forms but also search for occasions to use the L2 for communication. Promotion of out-of-class learning is a way to give students more autonomy about what they learn and how related to the topic of study, as well as a chance to reflect more deeply on their goals, methods of study and successes and failures. Benson (2013) describes how learners can actively and autonomously interact with material, social or technological resources available to them to direct their own learning process, which in turn will support their learning and give them the opportunity to reflect on their learning process. This study will examine whether any autonomy the students have in how they communicate, what they communicate about and how much during the VE is a connected factor to the learners' motivation to practice English outside of class.

Research has shown that communicating with first language (L1)-different

interlocutors allows for greater negotiation of meaning through interaction and chances for production (Bueno-Alastuey, 2013). Interaction is a dynamic mechanism for improving L2 competence (Mackey & Goo, 2007) because it creates more substantial opportunities to notice the target forms (Mackey, 2006) and gaps between these forms and the learners' own interlanguage before producing output (Gass & Mackey, 2006).

Learners using the L2 outside the classroom are using language authentically as a lingua franca and adapting it in experimental ways to achieve a communicative goal (Dewey & Jenkins, 2010). This investigation will explore whether the learners found the ability to communicate with peers around the world and learn about their culture intrinsically motivated them to use the L2 outside of class or whether the short-term instrumental goal of completing the task or a short interaction was their motivation.

There are also differences in learners' production based on whether they are speaking to a native speaker (NS) or non-native speaker (NNS). This study is based upon a VE between NNSs. Research has shown that communication with other NNSs results in greater modified output than exchanges with NS (Sato, 2006) and also increased time on task and L2 output produced (Bueno-Alastuey, 2013).

Block (2007) refers to the importance of building an identity related to the second language whilst it is being learned and states that this identity is shaped strongly by crucial encounters of using the language outside of the classroom. In the classroom, among their peers, students may already have a strong identity related to their use of their first language or wish to present one. Interaction in an EFL classroom of learners who share an L1 can be problematic as there is a reluctance to communicate in the target language amongst peers and a likelihood to regress to adoption of their L1 when a breakdown in communication occurs (Freiermuth & Jarrell, 2006).

The use of text based communicative activities can encourage learners who are reluctant to speak in class to produce output and this can increase their confidence to produce oral output in future tasks (Beauvois, 1995, as cited in Thomas, 2017, p. 201). Low ability learners through communication in text chat or by uploading images or videos related to a discussion can feel a sense of involvement and community within the class, learners can also create another identity in the L2 or online world, which could be more outgoing or confident (Lamy, 2006; Thomas, 2017, p. 202). Hyland (2004) after studying the use of English outside of class by university students in Hong Kong, found that the students felt more in control when communicating in the out of class setting and found it less threatening to their identities (both individually and as part of a group). This study will investigate whether the VE helps to cultivate an L2 identity in the learners and whether this increases motivation to practice the L2 out-of-class.

Computer mediated communication (CMC) has been shown in research to have many positive benefits as a setting for interaction to take place (e.g. Canals, 2020; Smith, 2004; Ziegler, 2016). Chun (1998) suggests CMC can often create a

reduced stress environment in comparison to face to face (FTF) interaction leading to more willingness to communicate.

CMC used in VEs can be both synchronous or asynchronous and take place between different types of interlocutors for example NS and NNS or NNS and NNS. One VE which is growing in popularity among teachers of EFL at university level is the International Virtual Exchange (IVE). IVE will be the subject of the research in this paper.

Method

Setting

IVE is an online exchange where students interact mostly asynchronously because with students being in different time zones around the world this avoids scheduling difficulties. However, if students are using the forum at the same time they can communicate synchronously. It is run using a website: iveproject.org which connects them to the Moodle platform (a popular learning management system), where they are put into message board style forums to discuss particular questions or topics. Exchanges are conducted over 8 weeks using English as a lingua franca among the members of the forum, who are from an array of different countries. Students can post and reply using text, audio or visual (picture or video) content created by themselves. IVE is run using funding from a Japanese government grant and assistance from Hosei University, so it is free for any educational institution's students to join.

The IVE goals for students are as follows:

- 1** To improve your intercultural competency.
- 2** To experience authentic communication with students from other cultures.
- 3** To find out about your own and others' cultures and lifestyles.
- 4** To improve your communication skills. In this project, this means learning to communicate in another language with people who do not know much about your culture.
- 5** To improve your digital literacy skills.

(IVE Project, 2022)

Participants

This study took place in the first semester of the 2021 school year at a large, prestigious, private university in eastern Japan. The institute is a co-educational facility consisting of 10 undergraduate colleges which include 27 departments. It is included in the Japanese government's Top Global University project allowing it to receive funding with the goal of promoting globalization and internationalizing Japanese higher education.

The research participants for this study were the 59 students of six first year Discussion classes. They were all 18–19 years old and Japanese. The Discussion course is designed to equip students with communication skills and increase their level of fluency while taking part in discussions in English based on a variety of

topics, such as Becoming Independent, The Globalization of Japanese Culture and Public Behavior. The course is very student centred with a communicative and collaborative focus. The course meets once a week for 90 minutes over the 14-week semester. Participants included thirty level 3 students (TOEIC level 285–500), twenty level 2 students (TOEIC level 500–700) and ten level 4 students (TOEIC level under 285). The place of implementation was the online Discussion classes attended by these students at the university and taught by the researcher and author of this paper. Classes are separated not only based on their English level but also by the faculty the students are members of. Two classes were the Faculty of Community Welfare, one was Tourism, one class was Literature, one was Law and one was Economics.

Instrumentation

This study took a mixed methods approach consisting of a mixed closed-ended and open-ended questionnaire. The questionnaire asked students about their experiences of using the IVE Project as part of the course as well as about prior experiences using online forums. The questionnaire was given as a Google Form and asked a set of questions to collect both quantitative and qualitative data regarding their experience of using online forums, their motivation for studying English and their feelings prior to using the IVE Project, during and after. The researcher wanted to ascertain how much experience they had of using online forums prior to participating in the IVE project as part of this class. Questions were also asked to determine if their reason for studying English had changed as a result of the experiences they had communicating with foreign students using English. These questions were open-ended in order to give the students as much freedom as possible to explain their reasons for study and they were allowed to state more than one reason. No options were given and students were free to write their own answers in the space provided. Other questions were asked to gauge how they felt about the experience of using English to communicate with and learn about people in other countries or areas. Again, these questions were open-ended and students could write anything in the space. This was done in order to gather as much data as possible about what the students liked and disliked about the experience. Finally, questions were devised to discover how useful they felt the VE was as a tool for improving their English and to measure any change in their level of motivation, as a result of the experience.

The procedure ran as follows. At the beginning of the semester the students were enrolled in the IVE Project, taught how to use it and told to make at least one post every week as part of their homework. They were told they could make more than one post and by making many posts and replying to other posts and interacting with other students their English was likely to improve. In the final class of the course, students were given the optional and anonymous questionnaire to complete. The research purpose was explained to them, and they were asked for consent.

Results

Prior to the IVE Project, few students had past-experience of using an online forum in their first language, which was Japanese for 100% of the participants in this study, even fewer students had used an online forum in English (see Table 1).

Table 1

Student online forum experiences prior to IVE Project participation

	With L1	Without L1	With L2	Without L2
N ^o of students	12	47	6	53
%	20%	80%	10%	90%

Table 1 shows that prior to the IVE Project 47 students (80%) had never used a forum in Japanese. 12 students (20%) had used a forum in Japanese. Table 1 also shows that prior to the IVE project 53 (90%) had never used an English language online forum before but 6 students (10%) did have some experience of using English on an online forum.

Table 2 shows significant changes in students' reasons for studying English after using the IVE Project.

Since the students gave open-ended, qualitative answers relating to their reasons for studying English, the researcher grouped answers together under representative categories in order to clearly display them in table 2. Students could give more than one reason, hence the answers number more than the total number of students. Looking at table 2, it is clear to see that the most popular reason for studying English pre-IVE Project was *to graduate university*. This answer was chosen by 14 students. This was closely followed in popularity by *to communicate/make friends with foreigners*, which was chosen by 13 students. The third most popular answer, chosen by 9 students, was *to pass an exam*. The fourth most popular, chosen by 8 students, was *to get a job*. Joint fifth, both chosen by six students were *to study abroad* and *to travel*.

After participating, 28 students said their reason to study English was *to communicate/make friends with foreigners*. The next most popular reason was *to study abroad* with 10 students, an increase of 4 students from prior to participation in the IVE project. Graduating university and getting a job are still in the top four most popular reasons with 8 and 6 students respectively stating these reasons, but these numbers have dropped from 14 and 8 students who gave these answers as their main reasons for studying English before taking part in the IVE project.

When the students were told that they would be participating in the IVE Project, many students felt it would be difficult and many others felt stressed or worried, however, this anxiety seemed to dissipate for many after they actually began using the VE (table 3).

Table 2

Student reasons for studying English pre- and post-IVE Project participation

Reason	Students pre	Students post	% pre	% post
To graduate university	14	8	24	14
To communicate/make friends with foreigners	13	28	22	47
To pass an exam	9	2	15	3
To get a job	8	6	14	10
For travel	6	3	10	5
To study abroad	6	10	10	17
For fun	4	2	7	3
To work abroad	4	5	7	8
Useful for the future	3	0	5	0
For homework	1	0	2	0
To improve communication skills	1	2	2	3
To live abroad	1	1	2	2
To read foreign news	1	1	2	2
To understand other cultures	0	1	0	8
To use IVE	0	5	0	2

Table 3

Students feelings about the IVE Project pre-participation and during (post-) participation

Feeling	Students pre	Students post	% pre	% post
Worried/stressed.	8	5	14	9
Thought it would be/was boring.	5	10	8	17
Thought it would be/was difficult.	36	25	61	42
Excited. Thought it would be/was fun.	10	19	17	32

Table 3 shows that the majority of students (61%) felt that using the IVE Project was going to be difficult before they started using it. 17% felt positive about using it, saying they were excited and expected it to be fun. 14% were worried or stressed about the thought of participating in the project and 8% thought that it would be boring. Table 3 also shows how the students felt while taking part in

the IVE project and using the online forum to make posts or read posts made by others. The results found that 42% found the VE difficult during the experience but 32% thought it was fun. 9% were worried or stressed while using the forum and 17% thought that it was boring.

The most popular topic IVE topic among the students was *the Cultures Around Us* with 46% followed by *Student Life in Our Cultures* with 27% (table 4).

Table 4

Most and least enjoyed topics

Topic	Students most enjoyed	Students least enjoyed	% most enjoyed	% least enjoyed
Who We Are	9	5	15	9
Student Life in Our Cultures	16	10	27	17
The Cultures Around Us	27	12	46	20
Gifts	7	32	12	54

Looking at table 4, *Who We Are* was chosen by 15% of the students and *Gifts* by 12%. The least popular topic was clearly *Gifts* as stated by 54% of the students. *The Cultures Around Us* was disliked by 20% of the students. *Student Life in Our Cultures* was the least enjoyable for 17% of the participants. *Who We Are* had the fewest number of students finding it their least favourite topic: 9%.

When making posts on the IVE Project most of the students were most concerned with comprehensibility (Table 5).

Table 5

The main thoughts of students when making a post on the IVE Project

Main thought	N° of students	%
Completing the post, so I could do something else	9	15
Writing a good post, so I could get a good grade	10	17
Writing a post that could be easily understood	28	48
Writing a post that would be interesting	12	20

Table 5 conveys that the most common thought when posting was a focus on making a post that could be clearly understood (48%), followed by a focus on making an interesting post (20%). 17% said they were thinking about writing a good post in order to get a good grade and 15% were thinking about just trying to complete the post so they could do something else.

The majority of the students found it interesting to read posts made by the other users of the forum (table 6).

Table 6

What students thought about reading posts by users from other countries/ parts of Japan?

Answer	N° of students	%
Not interesting at all	1	2
Somewhat interesting	7	12
Interesting	42	71
Extremely interesting	9	15

Table 6 displays that when reading posts made by others a large majority of 71% found them interesting and 15% found them extremely interesting. 12% said they found reading posts somewhat interesting and 2% not interesting at all.

The majority of students perceived the IVE Project as a useful tool for improving their English (68%) (table 7).

Table 7

How useful the students found IVE Project for improving their English?

Perceived usefulness	N° of students	%
Not useful at all	1	2
Somewhat useful	11	18
Useful	40	68
Extremely useful	7	12

Table 7 shows that 12% said it was extremely useful, 18% said it was somewhat useful and only 2% (1 student) said that it was not useful at all for improving their English.

The motivation level of students increased substantially between the pre-test and post-test (table 8).

Table 8

Level of motivation to study English before and after taking part in the IVE Project

Level of motivation	Students pre	Students post	% pre	% post
I had no interest at all in studying English	1	1	2	2
I didn't want to but I thought I need to so I will try	14	6	24	24
I wanted to study English a little	29	34	49	49
I was determined to learn English	15	18	25	25

Table 8 shows that before taking part in the IVE project, 15 students were determined to learn English, 29 wanted to study English a little, 14 didn't want to but thought they needed to try and 1 had no interest at all in studying English. After participating in the IVE project the determination to study English had increased substantially. 18 students said they are determined to study English, 34 said they want to study English a little, 6 said they don't want to but think they need to and 1 said he/she has no interest at all in studying English.

The questions about what students liked and disliked about participating in the IVE project were open ended, so a variety of answers were given. The researcher has tried to group answers which were common and significant to the research. When the students were asked what they enjoyed the most about using the IVE Project, 18 answers referred in some way to learning about/ sharing culture or communicating with foreign people. Answers included, "I was able to communicate with people from various countries" and "I enjoyed learning the different cultures". When they were asked what they did not like about participating in the IVE project 7 answers were related to technology, for example one student said, "The system was difficult". In addition, 7 answers were related to how many students were from the same country or from Japan. Here are two of the answers, "There were many Japanese people" and "Most people are Japanese so I wanted to communicate with people from other countries more".

Discussion

Despite an initial wariness of the project, most students adapted to and became appreciative of the advantages of participation in the VE. Students had little experience of using forums prior to the IVE Project. Table 1 shows 43 had never used a forum in Japanese and 53 had never used a forum in English, so that would explain their concern that it would be difficult and their stress or anxiety about using a VE (table 3). 7 students mentioned difficulty using the forum or the technology

as something they did not like about the VE. However, table 3 shows that most adapted to using the new technology quickly. Past research into the readiness of Japanese students to incorporate CALL into their curricula seems to show that they are far more proficient at using technology, such as their smartphones for gaming, entertainment or socializing than for any educational purposes (Lockley & Promnitz-Hayashi, 2012, as cited in Mehran et al., 2017). However, through learner training, awareness of the merits of tools such as VE and attitudes towards CALL can be positively influenced and this can lead to successful learning outcomes (Hubbard, 2005, in Mehran et al., 2017). This can be seen in the shift from negative preconceptions about using the VE to the change to more positive feelings while actually using the technology in table 3. Moreover, table 7 clearly shows that the majority of the students appreciated the technology as being useful for their English learning.

There seems to be focus placed on both form and meaning when making posts (table 9). Meskill and Anthony (2005) suggest that the extra time afforded to learners by text-based CMC allows for more time for processing and responding appropriately. Perhaps this is a factor in why the students found the exchange both interesting (table 6) and beneficial for improving their English (table 7).

Looking at table 2 the results show that for many, their motivation had changed from instrumental to more integrative (Gardner & Lambert, 1972). It seems to indicate that taking part in the IVE project created or nurtured a desire to meet and communicate with foreigners and experience their culture and to take their language learning outside of the classroom. The popularity of topics related to culture (table 4) seems to show that students greatly enjoy sharing their culture and learning about the culture of others using the forum.

The use of the IVE project outside of class seemed to increase intrinsic motivation in students, which according to Gardener (2010) is one of the attributes of an effective language learner. Table 8 seems to indicate that the IVE Project was able to increase the motivation of students to study English outside of the classroom. When looked at alongside the increase in intrinsic reasons for study (table 2) and the popularity of culture centric topics (table 4), it could be perceived that the use of English as a means to learn about other cultures and share one's own culture could have been a factor in the increase in motivation. Dörnyei & Csizér (2002) state that integrativeness as a temperament is a cognitive and emotive "*identification*", which Garden & Lambert (1972) say stems from being around native speakers of the L2 and their culture.

Block (2007) suggested that learners can create a different identity associated with the L2. Isabelli-García (2006, as cited in Block, 2007) claimed students studying abroad who integrated with the local community to practice the L2 developed a less ethnocentric and more ethnorelative and intercultural identity. Further in relation to identity, could the creation of an online persona in the VE have reduced feelings of anxiety associated with face to face communication among peers and fed a desire to conform and fit in with their online peers (Bullingham & Vasoncelos, 2013). Although further research would be required to answer these

questions other research purports the use of CMC for lowering learner anxiety and increasing confidence: Freiermuth and Jarrell (2006) found that CMC generates greater WTC because it provides a suspension of the social rules of face-to-face communication and they found some learners found this environment for communication preferential.

In relation to the complaint from seven students that the IVE Project had too many participants from Japan or from the same country, this is probably a result of the fact that this VE was first set up in Japan. The creator is trying to grow the project and promote it in a variety of countries and the number of participating countries increase every year. However, it is still most popular in Japan and it seems students can unfortunately be placed in a group that has several participants also based in Japan. Their complaint about this seems to support their integrative motivation to use the language to communicate with foreigners and learn about their culture along with the fact 18 students gave these answers as the thing they liked most about the VE.

Conclusion

The answer to RQ1 is that a virtual exchange using text-based communication can motivate students to use English outside of the classroom. The answer to RQ2 is difficult to ascertain and seems to depend on the individual learner. However, the experience of being able to communicate with people from other countries using the L2 seems for many students to increase their integrative motivation to study the language and so VE should be utilized as a tool to increase practice of the L2 outside of the classroom. Language acquisition requires frequent encounters and opportunities for analysis of the rules of correct usage over a long period of time (Ellis, 2002), as well as opportunities to test the understanding of these rules with output (Swain, 1985). Learners require attentive practice of the language both inside and outside of the classroom in order for acquisition to take place (Hyland, 2004; Nunan & Richards, 2015). Therefore, if CMC can increase the motivation of learners to use English outside of the classroom its use should be explored as a means of bolstering language learning. More research should be conducted involving the benefits of different types of virtual exchanges both synchronous and asynchronous, text-based and live video chat based. Moreover, further research into the effects of communicating with other NNS on L2 identity could also have pedagogical implications and should therefore be encouraged.

References

- Benson, P. (2013). *Teaching and researching: Autonomy in language learning*. Routledge.
- Block, D. (2007). *Second language identities*. New York: Continuum.

- Bueno-Alastuey, M. C. (2013). Interactional feedback in synchronous voice-based computer mediated communication: Effect of dyad. *System*, 41(3), 543–559.
- Bullingham, L., & Vasconcelos, A. C. (2013). 'The presentation of self in the online world': Goffman and the study of online identities. *Journal of information science*, 39(1), 101–112.
- Darhower, M. A. (2008). The role of linguistic affordances in telecollaborative chat. *Calico Journal*, 26(1), 48–69.
- Canals, L. (2020). The effects of virtual exchanges on oral skills and motivation. *Language Learning & Technology*, 24(3), 103–119.
- Chun, D. (1998). Signal analysis software for teaching discourse intonation. *Language Learning & Technology*, 2(1), 74–93.
- Dewey, M., & Jenkins, J. (2010). English as a lingua franca in the global context: Interconnectedness, variation and change. *Contending with globalization in world Englishes*, 72–92.
- Dörnyei, Z., Csizér, K., 2002. Some dynamics of language attitudes and motivation: results of a longitudinal nationwide survey. *Applied Linguistics* 23, 421–462.
- Ellis, N. (2002). 'Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition'. *Studies in Second Language Acquisition* 24:143–188.
- Ellis, R. (2008). *The study of second language acquisition*. 2nd ed. Oxford: Oxford University Press.
- Freiermuth, M., & Jarrell, D. (2006). Willingness to communicate: can online chat help? 1. *International journal of applied linguistics*, 16(2), 189–212.
- Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House Publishers.
- Gass, S., & Mackey, A. (2006). Input, Interaction and Output: An Overview. *AILA review* 19(1), 3–17.
- Hyland, F. (2004). Learning autonomously: Contextualizing out-of-Class English Language Learning. *Language Awareness*, 13(3), 180–202.
- IVE Project. (2022, June 16). What is the International Virtual Exchange Project? Retrieved from: <https://iveproject.org/mod/page/view.php?id=34#1>
- Jung, Y., Kim, Y., Lee, H., Cathey, R., Carver, J., & Skalicky, S. (2019). Learner perception of multimodal synchronous computer-mediated communication in foreign language classrooms. *Language teaching research*, 23(3), 287–309.
- Klein, W. (1991). 'Seven trivia of language acquisition' in L. Eubank (ed): *Point Counterpoint: Universal Grammar in Second Language*. Amsterdam, John Benjamins.
- Lamy, M. N. (2006). Interactive task design: Metachat and the whole learner. *Investigating tasks in formal language learning*, 242–264.

- Lightbown, P. M., & Spada, N. (2021). *How Languages Are Learned 5th Edition*. Oxford university press.
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied linguistics*, 27(3), 405–430.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis.
- Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS–NNS and NNS–NNS adult and child dyads. *Language learning*, 53(1), 35–66.
- Mehran, P., Alizadeh, M., Koguchi, I., & Takemura, H. (2017). Are Japanese digital natives ready for learning English online? A preliminary case study at Osaka University. *International Journal of Educational Technology in Higher Education*, 14(1), 1–17.
- Meskill, C., & Anthony, N. (2005). Foreign language learning with CMC: Forms of online instructional discourse in a hybrid Russian class. *System*, 33(1), 89–105.
- Nunan, D., & Richards, J. C. (Eds.). (2015). *Language learning beyond the classroom*. Routledge.
- Sato, M. (2006). Modified output of Japanese EFL learners: Variable effects of interlocutor vs. feedback types.
- Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in second language acquisition*, 26(3), 365–398.
- Swain, M. (1985). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development. In Gas & C. madden (Eds.) *Input in Second Language Acquisition*. Rowley, MA.: Newbury House.
- Terrell, S. S. (2013). Integrating online tools to motivate young English language learners to practice English outside the classroom. In *Explorations of Language Teaching and Learning with Computational Assistance* (pp. 184–192). IGI Global.
- Thomas, M. (2017). *Project-based language learning with technology: Learner collaboration in an EFL classroom in Japan*. Routledge.
- Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition*, 38(3), 553–586.

Appendix

Student questionnaire

I am conducting a research project into students' motivation to study English and the use of online forums to study English. I would appreciate if you could answer the questions honestly. None of your personal information will be shared and your answers will only be used for research purposes. The survey is completely optional and will in no way affect your grade.

- 1** Do you consent to your answers being used in this research project?
Yes. No.
- 2** Are you over 18?
Yes. No.
- 3** Before taking part in the IVE project, did you use any online forums in your first language (e.g. Japanese)?
Yes. No.
- 4** Before taking part in the IVE project did you use any online forums in English?
Yes. No.
- 5** Before taking part in the IVE Project, what was your reason for studying English?
- 6** Before taking part in the IVE Project how motivated were you to study English?
 - I had no interest at all in studying English.
 - I didn't want to study English but I thought I need to so I will try.
 - I wanted to study English a little.
 - I was determined to learn English.
- 7** How did you feel when you were told you were going to use an online forum as part of this English class?
 - I was worried/stressed.
 - I thought it would be boring.
 - I thought it would be difficult.
 - I was excited/ I thought it would be fun.
- 8** How did you feel when you were using the online forum?
 - I was worried/stressed.
 - I thought it was boring.
 - I thought it was difficult.
 - I thought it was fun.
- 9** Which topic did you enjoy the most from the IVE project?
 - Who we are
 - Student life in our cultures
 - The cultures around us
 - Gifts
- 10** Which topic did you enjoy the least from the IVE project?
 - Who we are
 - Student life in our cultures
 - The cultures around us
 - Gifts

- 11** When making a post on IVE project what were you thinking about the most?
- Completing the post so I could do something else.
 - Writing a good post so I could get a good grade.
 - Writing a post that would be easily understood by other students using the IVE project
 - Writing a post that students from other countries parts of Japan would find interesting.
- 12** How interesting did you find reading posts by students from other countries or parts of Japan?
- Not interesting at all.
 - Somewhat interesting
 - Interesting
 - Extremely interesting.
- 13** After taking part in the IVE Project, what is your reason for studying English?
- 14** After using the IVE project, forum how motivated are you to study English now?
- I have no interest at all in studying English.
 - I don't want to study English but I think I need to so I will try.
 - I want to study English a little.
 - I am determined to learn English.
- 15** How useful do you think using the IVE Project forum was for improving your English?
- Not useful at all.
 - Somewhat useful.
 - Useful.
 - Extremely useful.
- 16** What did you enjoy the most about using the IVE Project online forum?
- 17** What didn't you enjoy about using the IVE Project online forum?

Author's bio

Jack Pudelek has taught English in Japan for over 12 years. He has an MEd in TESOL and Applied Linguistics. His recent areas of interest for research include CALL, CLIL and learner goal setting.

3

Self-access learning and Minecraft: Observations and student perceptions

Robert Remmerswaal and Robert Dykes, Sojo University

Abstract

Minecraft is an interactive virtual sandbox game that came out in 2011. The authors created a Minecraft server for their students at a private Japanese university to expand the offerings of their Self-Access Learning Center. While Minecraft has been shown to promote communication and collaboration in a structured course, it is unknown if this can be replicated in a less structured scenario. Over the course of a year, students were able to access the server anytime from anywhere. All students who used the server were asked to use an “English Please” mentality. English was not strictly enforced but instead encouraged with a focus on communicative output rather than grammatical competence. The authors designed several events using the community of inquiry framework. Quantitative data was collected in the form of a student survey, recorded event attendance, and server log records. Qualitative data was collected through an open-ended survey, teacher notes and observations, and server chat records. From this data, many students appear to be interested in Minecraft, but only a few were interested in playing Minecraft while using English. Students who used the server regularly appeared to improve their English communication skills. The educational value of the different events is discussed, with group-based events believed to be the most valuable for building a community.

Keywords: Minecraft, self-access learning, community of inquiry

Introduction

In August 2021, the authors launched a Minecraft server in Sojo University’s Self-Access Learning Center (SALC). Minecraft was positioned as a tool to be used in a social constructivist approach to learning English. In social constructivism,

meaning is constructed through interactions with others (Kalina & Powell, 2009). In line with second language acquisition concepts expressed by Firth and Wagner (1997), the authors stressed communicative competence rather than grammatical competence for language to be learned through social interactions while playing Minecraft. The rationale for choosing Minecraft, technical issues, and some of the activities used in this study have been described in an earlier paper by Remmerswaal (2022).

Minecraft is an open sandbox game, meaning that rather than a predetermined goal or linear gameplay style, it is a space that allows for creativity. Within the game, blocks can be arranged and structured to build or mimic anything from the real world (Nebel et al., 2017). It is the best-selling game of all time, with 238 million units sold (Microsoft, 2021). Minecraft was released in 2011; however, in the past five years, player numbers have tripled from 40 million to over 140 million (Statista, 2021). The participatory culture of the game allows for creativity and communication (Stevens, 2021). Playing Minecraft collaboratively elicits the language needed for problem-solving, creativity, and collaboration (Chien, 2019). While there are many collaborative games available, Minecraft allows the teachers to create their own Minecraft server easily, thus retaining the ability to moderate users, content, and tasks, ensuring a safe place for students. It was these factors, along with the low barrier to entry, that led to the decision to introduce the Minecraft server in the SALC (Remmerswaal, 2022).

This exploratory paper looks at the first year of Minecraft in the SALC. Student participation in the Minecraft server was voluntary, with no commitments necessary to play. Students were encouraged to participate in various events, some face-to-face and others with remote participation, with the server always open for unstructured play. A crucial part of online play is a community, as it gives a purpose for being in-game and allows students to learn and improve from each other (Stevens, 2021). This study looks at the way students and teachers use Minecraft through the lens of a Community of Inquiry (CoI) (Garrison & Arbaugh, 2007). The research questions are:

- 1 Are Sojo University students interested in playing Minecraft while using English?
- 2 What Minecraft activities are conducive of an education experience, as defined by CoI, within a SALC?

Literature review

Gaming as a learning tool has the ability to create fully engaged participants who are proactive, collaborative, reflective, critical, creative, and innovative problem solvers (Gee, 2013). Gaming promotes a participatory culture with systematic thinking and experimentation (Squire, 2011). Learning in games can be made socially meaningful, a notion put forth by Vygotsky (1986), as learners can interact with peers, teachers, and experts. Games allow students to immerse themselves

and experience the environment rather than simple written or visual descriptors (Bogost, 2011; Nguyen, 2016).

Minecraft has appeared in a number of peer reviewed papers for learning in a variety of contexts (Alawajee & Delafield-Butt, 2021; Baek, et al., 2020). There are endless possibilities of what can be built in Minecraft. This allows students to engage in their own way and leads to distinct outcomes for each individual (Marcon, 2013). Several studies have demonstrated that Minecraft can increase motivation, creativity, as well as communication and collaboration skills (Al-Washmi et al., 2014; Lorence, 2015; Pusey & Pusey, 2015; Uusi-Mäkelä, 2015; Callaghan, 2016).

In the context of language learning, a few studies found Minecraft allowed students to explore scenes from different stories and discuss and reflect on the content (Cipollone et al., 2014; Marlatt, 2018). One study used the characters and scenes from the game as inspiration for student stories and blog entries (Marcon, 2013). Others had students write journal entries or short stories in the game, for others to find and interact with (Uusi-Mäkelä, 2014; Lorence, 2015). Minecraft as inspiration for writing was summarized by Kuhn and Stevens (2017) to be ideal for language learning as the open and dynamic game makes the experience enjoyable and leads students to develop composition skills to describe their game experiences. There was an additional case of children becoming fluent in English through watching Minecraft videos, playing with a community, and then making their own videos to share (Smolčec et al., 2014).

Minecraft, when played on a server, is social. The experience gained by a learner through trial and error leading to the final product can be shared both in and out of the game (Cipollone, 2014). One study had students use written exchanges in the target language and found that success in the given tasks relied on the sharing of knowledge and resources (Uusi-Mäkelä, 2015). Another found that increased spoken communication facilitated collaboration between participants sharing expertise, creating objects, solving problems, or completing a building task; and improved teacher-student relationships (Callaghan, 2016). Participants reported enjoying the activities and improving their English through the collaboration needed for task completion (Swier, 2014). Similarly, Egbert and Borysenko (2019) found Minecraft was a prompt for discussions with 73% of the pre- and -in service language teachers from the study wishing to use Minecraft in their own language classes. In one case, first-time female users of Minecraft improved their collaboration and communication skills as they used expressive language when asking for assistance (Marcon & Faulkner, 2016). Both Petry (2018) and Nebel et al. (2017) found similar findings that collaboration led to increased learning and interaction.

Beyond language learning, Minecraft was shown to promote information literacy skills (Bebbington & Vellino, 2015) and digital literacy (Herold, 2015). However, there are a few areas of concern for teachers. One is it is easy for students to be off task, exploring or building something not connected to a goal. While some report Minecraft as being easy to learn (Smolčec et al., 2014), some pre-service teachers in McColgan et al. (2018) disagreed. They found the learning curve quite

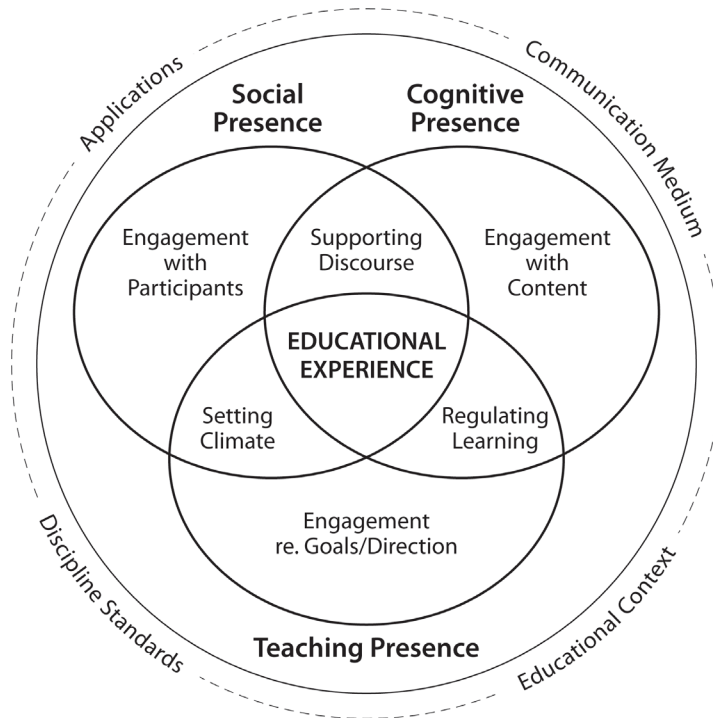
steep for themselves as teachers, that tasks required a lot of time to design, and worried students may find the game too complex. Dykes (2021) remarked on a similar finding in his broad overview of using commercial off the shelf games in language learning, that using games like Minecraft to supplement English learning can be a time-consuming endeavor.

Despite these potential challenges, Alawagee and Delafield-Butt (2021) concluded that Minecraft can be a mediator between players and academic content in teaching first or second languages, in a well-designed lesson. In all the reviewed literature, Minecraft was either used in a classroom or for a set period of time, with specific students, in order to complete specific tasks. The researchers were unable to find an example of Minecraft used in an unstructured manner. This is why an exploratory study of Minecraft within a SALC is of interest. It may introduce new findings in the educational role Minecraft has when participants join as frequently or infrequently as they wish and when planned tasks must accommodate an unknown number of new and returning players.

Methods

Minecraft was available at any time on or off campus, with a purposeful educational experience for students facilitated through a variety of online and in-person activities for individuals and groups. This was organized through the framework of Garrison's Community of Inquiry (CoI). CoI is a model of the essential elements needed for a successful online higher education learning experience, which has also been used for hybrid learning (Drysdale et al., 2013). The goal of a CoI is to create a deep, meaningful, collaborative learning experience by developing the social, cognitive, and teaching presences. As can be seen in Figure 1, social, cognitive, and teaching presences overlap in CoI to create the educational experience (Garrison & Arbaugh, 2007).

Figure 1
Community of Inquiry Framework



Note. From *The Community of Inquiry*, n.d. CC-BY 4.0

Social presence requires open communication, group cohesion, and effective expression. Students should be able to identify with the community in a trusting environment and develop their relationships while showing their personalities (Garrison, 2009). When collaboration and community are the goals, initial activities should center on exploring and negotiating expectations to create a climate for open communication (Garrison, 2009). Social presence was facilitated with in-person events, casual conversations in the conversation lounge, with the on-line social presence utilizing Teams, Discord, and in-game chat. Another benefit of Minecraft is the use of an avatar, which has been shown to increase collaboration and task performance compared to a non-avatar condition (Pan & Steed, 2017).

Cognitive presence requires a triggering event, exploration, integration, and a resolution. Some indications of these are information exchanges, connecting ideas, or the application of new ideas (Garrison & Arbaugh, 2007). These triggering events were the cognitive processes required to complete the tasks designed by the teachers. This could be searching for the correct English words to hold a short conversation, the words needed to collaborate, finding others to complete a task, or finding other ways to collaborate.

Teaching presence includes the design and organization of the class, the facilitation of discourse, and direct instruction. Some indicators of this presence

include shaping constructive exchanges and focusing on and resolving issues (Garrison & Arbaugh, 2007). Tasks were all designed for specific purposes. As part of the teaching presence, the authors actively participated in the Minecraft server and created events throughout the year. This allowed the authors to make changes and adapt the events and ways of communication as the year progressed based on student feedback and observations.

Data collection

To better answer the research questions, several forms of data were collected. These methods are summarized in Table 1.

Table 1

Methods used to answer the research questions

Research question	Data collection methods
Are Sojo University students interested in playing Minecraft while using English?	Survey 1
	Event attendance
	Server records
What Minecraft activities are conducive of an education experience, as defined by CoI, within a SALC?	Teacher notes and observations
	Server chat records
	Survey 2

Before investing in a server, Survey 1 was administered to students in late July 2021 and contained four questions related to Minecraft. These questions were asked in Japanese and trialed with two teachers to make sure responses were appropriate. The questions were designed specifically to understand if students had an interest in a SALC Minecraft server.

- 1) Do you play Minecraft?
- 2) At the SALC, if we use Minecraft to have fun and practice English a little bit, would you join a SILC Minecraft server? (You can build and compete with other students online).
- 3) What do you use to play Minecraft? (Choose one or more) (only available if the answer to question 1 was yes).
- 4) Would you like to receive emails with more information about SILC's Minecraft server?

The researchers were solely focused on Minecraft because of the flexibility of a server, the ability of the game to be cross played on any server, and its ability to circumvent some of the networking restrictions on campus. The survey was made to determine if students would be open to playing Minecraft collaboratively.

Server records and event attendance were then put in place to record the

activity of students. It is one thing to be interested in Minecraft and another to play Minecraft. While a survey can justify an initial commitment, actual use of the server is needed to demonstrate a true interest in the platform. The usernames of any person who joined the server was registered in the server (unique users), allowing the researchers to track the total number of unique users to join the server. Usernames are unique to each paid account and can only be changed for a fee paid to Microsoft.

The researchers primarily used observations to answer the second question, with Survey 2 and server chat records used as supplemental sources. The authors kept field notes during planned events and impromptu conversations with students in a manner detailed by Richards (2003). Impromptu conversations happened in both English and Japanese but were summarized in English. These summaries were typed into a Microsoft Form which included the date of the interaction, the site, number of people, notes on speaking patterns, types of student interactions (questions, comments, responses), room for other descriptions, and an area to suggest improvements. In reference to these notes, the two authors also completed a written reflection in June 2022. Both authors read the notes of the other and discussed any differences in what was observed at an event.

A large part of the educational experience relied on communication. From January 23, 2022, a plugin was added to record the chat history of students in the server. Prior to that time, a teacher was required to observe any chats in real time while playing Minecraft. This was not a reliable metric. With the addition of the plugin, all chats have been recorded since that date, including the username and time of the chat. This was used to understand the types of conversation during unstructured play.

The second survey, in January 2022, followed up with the respondents of the first survey via the emails they provided and included any other student who had joined the Teams group. Originally 130 students expressed interest in the Minecraft server, but most never took part in any activities. Other students only participated once. The researchers wished to understand some of the reasons students had for not attending, or their limited attending. The authors included five open questions and one multiple choice question. It was unclear at the time if the lack of participation was connected to Minecraft, the SALC, or the sharing of event information. These questions were asked in Japanese.

1) If you did not participate, why not? If you did participate, what did you enjoy and what would you like to see changed or added?

2) If you know Minecraft, which game mode do you like best? (Creative, Survival, PvP, Contest).

3) Information about SILC Minecraft was sent via Teams. Would you prefer to get information through other means, such as newsletters or email?

4) Have you participated in any SALC activities? (Please select all that apply)

5) May we contact you if additional questions arise? (If available, please provide email)

6) Do you have any other comments or suggestions?

Data analysis

Surveys 1 and 2 were administered with Microsoft Forms and analyzed within Microsoft Excel. Quantitative data was analyzed by category totals. Qualitative data was coded by both researchers synchronously with any differences in opinion negotiated between them. Researcher observations and reflections were reviewed and discussed by both researchers before their inclusion in this paper. Server data was exported from Discord into Microsoft Excel. Student conversations were reviewed and counted by both researchers, with no differences to negotiate. The chats included student-student (S-S), student-teacher (S-T), and teacher-teacher interactions. The researchers limited the chat count to S-S and S-T interactions that occurred in English.

Participants

English is a required course for two years at Sojo university. There are no English majors, and the average student English level is A2 on the CEFR scale. The SALC is available for all undergraduate and graduate students but is most frequented by first- and second-year students. The Minecraft server was aimed at these first- and second-year students as they attend English classes in the SALC building but was advertised throughout campus for all students. Both researchers were also participants in the events. From late April, Taiwanese students from the National Taiwan University of Education (NTUE) were invited to the server. These students were all studying to be English teachers.

Researcher positionality

The researchers both had experience playing Minecraft collaboratively with fond memories. One was particularly interested in Minecraft for students after witnessing his nephew improve his English as they played Minecraft together (Remmerswaal, 2022). Both researchers were interested in having more students participate in the SALC and believed Minecraft had the potential to draw-in students who may not otherwise utilize the SALC. This experience with Minecraft influenced the types and design of activities. The researcher's positive views of Minecraft is also a potential bias in the interpretations of what they observed throughout the year.

Minecraft event design procedures

Several types of events and modes of play were available to the students, including solo play, local group play, planned group play, and events (Remmerswaal, 2022). All the activities encouraged players to use the resources available to them, such as other students, teachers, or their smartphones to cooperate in English. These events were designed in consideration of the social, cognitive, and teacher presences for the CoI model. Table 1 describes these in more detail. It must be noted that the events were never designed more than two at a time. Typically, they were designed based on the results of the previous events. Table 1 represents what was planned whereas Table 2 describes what was observed.

Table 1

Description of each Minecraft event and the intended educational experience

Event description	Educational experience
<p>Pre-launch event (September). Send the details to interested students. Let students explore the newly created world without instructions.</p>	<p>Social Presence: Students can build structures as a way to showcase personality and to spark future conversations. In-game chat is available. Teachers are accessible via email and Teams.</p> <p>Cognitive Presence: Students explore the world and find their own area to build structures. Students can make their structures before the official opening.</p> <p>Teacher Presence: Teachers can build their presence and interact with students or with the structures students build.</p>
<p>Halloween Build Contest (October). In creative mode, students build a Halloween themed structure. Each person uses a 50x50 square as a base.</p>	<p>Social Presence: In creative mode students are free to build without the need to gather blocks. This allows a showcase of personality. Students can discuss their builds in the conversation lounge, using in-game chat, and through Teams.</p> <p>Cognitive Presence: Using the area given to them, students decide the best way to use their space. They can learn from what others are doing or from online searches of Halloween builds. Hopefully they will become more comfortable in the server community.</p> <p>Teacher Presence: The teacher must create the designated areas either manually or with computer commands. The teacher should create an example for the students. The event and login details need to be available online and within the SALC with advertising wherever possible.</p>
<p>Minecraft Wednesdays (November and 2 weeks in December) Meet on campus with an online option. Play Player vs. Player (PvP) games. Students are put in teams or pitted against each other.</p>	<p>Social Presence: Students strategize in teams face-to-face or using Teams. Students must share their understanding of the game and their strategies to win.</p> <p>Cognitive Presence: Students are put into this arena and must figure out the rules of the game, a strategy to win, and how to utilize their teammates to defeat the map. Students quickly realize whether they truly understood instructions based on the outcome of their character movements.</p> <p>Teacher Presence: Teachers must find or build appropriate maps that can support different game modes and can be swapped in and out on the server. Students must demonstrate the appropriate vocabulary to work cooperatively and give assistance to students as they struggle to share meaning.</p>

Let's Build Saturdays
(November) Meet online to play Minecraft together in survival mode. Players can work on their own projects, on a shared project, or explore together.

Social Presence: Discussion is supported through Teams. Students are encouraged to assist one another, make suggestions, and build out the shared community that has formed. Students are able to take leadership roles and to demonstrate their skills or curiosity.

Cognitive Presence: Students interact with their environment, either exploring or building what is of interest to them. Students can choose to follow other players or to stay in their own area.

Teacher Presence: The teacher must have some building projects in mind or an area to explore. The teacher can ask about what students are doing and where they want to go in order to spark conversations.

Christmas Build Contest
(December)
Similar to the Halloween event, students build Christmas structures on a 100x100 square area.

Social Presence: This is a second opportunity for students to showcase their personalities and learn about other students through what they build. With a stronger understanding of the other members, there is increased opportunity to speak with other members about their builds using in-game chat, Teams, or in the conversation lounge.

Cognitive Presence: Using the area given to them, students decide the best way to use their space. They can learn from what others are doing in their spaces or from online searches of Christmas builds.

Teacher Presence: The teacher must create the designated areas either manually or with computer commands. The teacher should create an example for the students. The event and login details need to be available online and within the SALC with advertising wherever possible.

Unstructured Play (upgraded in January, but always available)
The server is expanded to allow multiple game modes and multiple maps to run concurrently. Students are encouraged to host their own events or meet on the server with friends.

Social Presence: Students continue to have access to in-game chat and Teams for their own chats, as well as face-to-face meetings they plan. Students are encouraged to plan events in the Teams chat so others can join. Students can now interact with the teacher asynchronously through written or video reflections.

Cognitive Presence: Students continue to interact with the environment and work collaboratively on their own schedules. Students are able to reflect on their journeys as well as practice using English using written and/or video journals.

Teacher Presence: To allow more creativity and exploration, multiple worlds are combined using a plug-in. Students can travel through in-game gates to switch between mini-games, creative mode, and survival mode. Students are given freedom to create their own events, work alone, or play with friends. A journal tool and a video log tool are also given to encourage reflection of learning.

Build-a-bed
(April) Start the new school year with a contest. Students build a bed on the server, take a screenshot and fill out a short form for a chance to win a gift card.

Social Presence: Students are incentivised to join an existing community without showcasing their own skill. New students can explore and see what exists in the server world before demonstrating their own skill.

Cognitive Presence: Students join the server to complete a simple task and enter a contest. From there, they can explore and learn about the other students in the community simply by exploring what has already been built.

Teacher Presence: The teacher must create a form and a how-to page for students to upload their information. The teacher must be available for questions and advertise the event.

Protect your Village (May, single day) A collaborative event held between Sojo University and the NTUE. Students work together in groups to defend a village. Each group is teleported to an existing village and builds defenses against mobs as well as new buildings to improve aesthetics, all while in survival mode.

Social Presence: Students communicate through discord with their teammates. Students share their ideas on how to protect the village or to make it more aesthetically appealing.

Cognitive Presence: Students use the blocks they have accumulated and search for new blocks to build and defend their village. They must share the resources they have and plan the collection of new resources to accomplish their building goals.

Teacher Presence: The teacher must create portals to villages throughout the world. The teacher must create a form to gather the information of interested students and divide those students into teams. The teacher creates a how-to guide and explains the contest at the beginning of the event. The teacher must rotate through the different groups to spark conversation, troubleshoot any issues, and offer support for the challenge.

Minecraft Wednesdays
(June and July)
Meet on campus at a time most players are available. Play in survival mode and work together to build and explore.

Social Presence: Students communicate with others in the room. Students can teach each other, work together, and interact with the teachers.

Cognitive Presence: Students interact with their environment, either exploring or building what is of interest to them. Students can choose to follow other players or to stay in their own area.

Teacher Presence: The teacher must have some building projects in mind or an area to explore. The teacher can ask about what students are doing and where they want to go in order to spark conversations.

<p>Destroy the Warden (July, single day) Students from both Sojo and NTUE work together to defeat the warden mob. Groups teleport to different warden areas. The first team to kill the warden is the winner.</p>	<p>Social Presence: Students communicate through discord with their teammates. Students share their ideas on how to defeat the warden in the most efficient way including how to share the resources they each have.</p> <p>Cognitive Presence: Students use the blocks they have accumulated and search for new blocks to kill the warden. They must share the resources they have and plan the collection of new resources to accomplish their goal.</p> <p>Teacher Presence: The teacher must create portals to different warden sites. The teacher must create a form to gather the information of interested students and divide those students into teams. The teacher creates a how-to guide and explains the contest at the beginning of the event. The teacher must rotate through the different groups to spark conversation, troubleshoot any issues, and offer support for the challenge.</p>
--	--

Ethical considerations



Both surveys required explicit consent from participants before they could answer questions. Students were made aware that chats were recorded in the Teams and Discord channels, as a welcome message within Minecraft, and on the how-to website. Any personal information that was collected has been anonymized for this publication. Institutional ethical approval was received from Sojo University.

Results

The researchers provided all English teachers of first- and second-year students with a link to survey 1 to give to their students. The possible pool of respondents was roughly 1600 students, but not all instructors made the survey link available to their classes. This was likely due to other surveys teachers often administer at that time for their research and a desire not to cause survey fatigue. The initial survey demonstrated some interest in Minecraft. From the 245 responses, 130 students indicated an interest in Minecraft for English, and 63 were active Minecraft players. With this information, the researchers proceeded to launch a server. While 130 students were interested in a Minecraft server, event attendance shows that fewer students would join Minecraft during planned activities. Table 2 describes the teacher observations of each event along with the total attendees (n=students unless otherwise indicated, N=total players including teachers). There are several students whose observations of their interactions were remarked upon. Aliases will be used: Yoshi, Kaneda, Tetsuo, Akira, Yuki, and Ken. All observations are in reference to the summarized field notes, reflections, and discussions of the two researchers.

Table 2

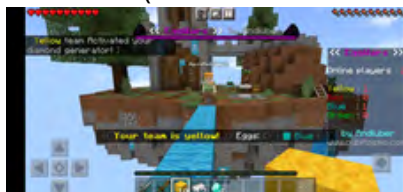
Summary of observations and the educational experience of each Minecraft event

Event observations	Educational experience observations
<p>Pre-launch event (September) One student created an incredible base that took many, many hours to compete. Another student created a giant hole. A third student joined, but a structure was not found. n=3, N=5</p> 	<p>Social Presence: One student communicated via in-game signs with the teacher. A second student sent a request via Teams for the teacher to increase the world difficulty. The third had no visible presence on the server. No apparent S-S interaction.</p> <p>Cognitive Presence: The two students explored the world and built structures. Students were able to make their structures before the official opening</p> <p>Teacher Presence: One teacher built a home and created some areas for students to explore and collect swords and armor.</p>
<p>Halloween build contest (October) In creative mode, students built a Halloween themed structure. Each person used a 50x50 square as a base. Each student had a very different structure. Teachers scored each structure and chose a winner, seen in the picture below. n=5, N=8</p> 	<p>Social Presence: Three of the students who participated spoke about the contest in the conversation lounge and asked questions. Two participated online only. There was no apparent S-S interaction.</p> <p>Cognitive Presence: The students used their areas well and created some amazing buildings. The three students who spoke with the teacher gained familiarity with the teacher, but other community bonds may not have been formed.</p> <p>Teacher Presence: The teachers created their own buildings for the students to view when they first arrived. Instructions were posted within Teams and posters with QR codes put up throughout campus.</p>

Minecraft Wednesdays

(November and 2 weeks in December) Two students were somewhat regular attendees. Four students visited one time only. Yoshi was a regular member in November, May, June, and July. In December, students Kaneda and Tetsuo were regular attendees. Other members typically attended only once. When Akira attended his first and only Minecraft Wednesday, he expressed how shocked he was that it was a teacher-led event (personal communication, November 17, 2021).

n=0-4 N=2-6 (min-max attendance)



Social Presence: Students used limited speech to coordinate their efforts, often using sounds to express dying in game, or being hit. Students could answer teacher questions and follow instructions to some extent. Students attempted S-S verbal cooperation with some success.

Cognitive Presence: Students demonstrated a knowledge of the rules and the ability to play the minigames. Cooperative ability grew over time. Student vocabulary did not clearly grow, but the ability to speak English while playing increased.

Teacher Presence: The teachers found several maps to play. Most worked well, but some had problems. Each map change required a server restart. Teachers were necessary to encourage interactions and elicit speech from students. Teachers provided vocabulary support on a few occasions within each active session.

Let's Build Saturdays

(November) Students met with a teacher on Saturdays for one hour of survival mode play. Yoshi and Kaneda were regular attendees. Conversations were held through MS Teams and were usually teacher-led. At the end of this event series, both students expressed interest in continuing to meet on Saturdays. One author explained that students could continue to use MS Teams and play together on Saturdays; the MS Teams channel was open for students to use it. However, both interested students stated they were uncomfortable starting a call independently (personal communication, December 3, 2021).

n=2, N=3



Social Presence: The teacher initiated the Teams call and initial conversation. One student would only speak when spoken to. The second student would ask questions and give some ideas of what they wished to do. Conversations were not complex, but understandable.

Cognitive Presence: Students and the teacher built out their own homes, explored the Nether area, and visited an ocean monument. S-T cooperation strengthened and vocabulary during discussions became slightly more complex with time.

Teacher Presence: The teacher suggested building a shared home structure, but each student wanted their own structure. The teacher often visited the two areas and offered resources or negotiated a sharing of resources between students.

Christmas Build Contest

(December) Similar to the Halloween event, students built Christmas structures on a larger 100x100 square area. Students occasionally sat together to build in Minecraft but were very focused on their own projects. Teachers scored each structure and chose a winner, seen in the picture below. While Kaneda was an active participant in the Halloween event, he made it clear that he was uninterested in another creative mode event (personal communication, November 25, 2021). In fact, only two participants from the Halloween event took part in this event.
n=6, N=9



Social Presence: The personalities of the students could shine through in this project. Students were able to share what they were building with the teacher with some detail. Very little S-S communication.

Cognitive Presence: From a creative standpoint, students used their spaces very well. There were snowmen, griches, a sleigh with a reindeer, a temple, and a building with moving walls. Students interacted with each other's structures, if not with each other.

Teacher Presence: The space was created using a flat map and commands to layer 100x100 areas with snow. Both teachers created their own Christmas themed buildings. Like all events, posters were put around campus with a QR code and details in Teams. The teachers engaged in many conversations with the players about what they were doing. Often while playing, but other times conversing without playing the game.

Unstructured Play (Changes from January) Unstructured play was available in September and November as well, but creative mode contests prevented world exploration in October and December. From January, a survival mode world and a creative mode world were simultaneously available. Students were encouraged to host their own events or meet on the server at will. One event was proposed by a student, where students and teachers built a railway over a Teams call. Students built their own structures independently during this time as well.



Social Presence: Students used in-game chat to speak with each other, occasionally at first, but with more frequency in May, once Taiwanese students were joining the server regularly. The one train station event had two teachers and two students collaborate on a train station design, the places the tracks could go, and negotiate who would provide which resources for the station. The new journal tool was used twice by one student. The video journal was never used.

Cognitive Presence: After greetings, resources within the game were the main cause for communication between students. Students had their own building goals and typically asked for specific blocks to accomplish them. There was no S-S communication related to working together to build a structure.

Teacher Presence: The teacher added a plugin that allows multiple worlds to exist simultaneously. Another plugin was added to record in-game chats. The teachers created a website to host all how-to guides and server information. The website included a journal tool and a video log tool to encourage reflection of what students were doing in the world and to give students a chance to use English asynchronously.

Build-a-bed (April) Four first-year students joined this contest and one student who had participated in previous events also joined. The four new students joined the protect your village event as well. n=5, N=5



Social Presence: New students did not interact with other students through in-game chat. This was likely due to the timing of joining the server when no other players were active.

Cognitive Presence: Students explored and learned about the other students in the community to some extent, seen by their beds being built near other structures.

Teacher Presence: The teachers created a form and a how-to page for students to upload their information. The teacher advertised the event throughout campus.

Protect your Village The teachers planned a collaborative event between Sojo University and the NTUE. Students worked together in groups of 5-6 to defend a village. Each group teleported to an existing village and built defenses against mobs as well as new buildings to improve aesthetics, all while in survival mode. Teachers scored each village after one hour and declared a winner. When asked about the experience, three students all reported that the event was enjoyable, but it was challenging to communicate with their team. Yuki explained he was hesitant to initiate the conversation, and the event ended as he started to feel comfortable (personal communication, May 6, 2022). During the conversation with Yuki, Ken interjected his shock that his friend had joined the event. He stated he would not be comfortable using English to play Minecraft (personal communication, May 6, 2022).
Japanese students n=7
Taiwanese students n=14
Teachers n=3



Social Presence: Students communicated through Discord with their teammates. Students shared their ideas on how to protect the village to some extent, with some students simply building on their own. With teacher intervention, teams all increased their communication and began to work more cooperatively.

Cognitive Presence: Students used their blocks to build walls and towers within the village. The sharing and categorization of blocks was the main point of communication.

Teacher Presence: The teachers created portals to the villages, though a glitch created some problems. Students were assigned to have a mix of Japanese and Taiwanese students. Teachers rotated through the different group chats to give advice, encourage group cooperation, and to assist, when possible, often by giving blocks.

Minecraft Wednesdays (June and July) Students in the Minecraft Teams channel were asked for their availability, with Wednesday afternoons having the most overlap. One student was a regular attendee, a second came twice. Five students joined once.
n=1-3, N=3-5



Social Presence: Students communicated with each other in mostly English, with some Japanese support. Students taught each other how to play. Students responded to teachers and described what they were creating or asked for assistance.

Cognitive Presence: The regular student used his resources to assist new students in creating their own structures or homes on the server. There was some cooperation with building. A teacher or student would give a tour of the community structures for new students to see and explore for themselves.

Teacher Presence: The teachers planned activities such as a railway expansion, exploring a Nether Fortress, and finding a luscious biome. Teachers often initiated conversations by inquiring on what students were building and offering assistance.

Destroy the Warden (July) Students from both Sojo and NTUE were invited. The date was near exams in Japan and a month into the Taiwanese summer vacation.
n=1, N=3



Social Presence: The student joined through Discord, but was unable to speak, but could listen. In-game chat then allowed for responses.

Cognitive Presence: The players all experimented with different ways of killing the warden. This included swords, TNT, fireworks, axes, and fire.

Teacher Presence: The teachers created portals to the wardens. When there were no teams, the teachers created a new challenge of killing multiple wardens. Teachers initiated most conversations, but this was a technology issue as well.

Note. All pictures were taken as screenshots during gameplay. n represents the number of students participating. N represents the total number of participants.

Beyond the attendance of the events, students were able to join the server for unstructured play. The server record of unique players along with the server record of chat histories is displayed in Table 3. Unique players include teachers and the five SALC devices with unique player IDs. Multiple students shared these devices. A few players may be represented twice if they have different identifiers on different devices. There were a number of events that took place between Sojo University and the NTUE. The only data collected on the Taiwanese students were the number of participants and the number of chat interactions when a Japanese student was in the conversation. The final two rows refer to messages sent using

the in-game chat function. Days with an in-game chat refers to any day with at least one English message sent within Minecraft. Total exchanges via in-game chat refer to the number of English messages sent with at least one Japanese student participating in the conversation. There were several cases of students sending messages in Japanese or Mandarin, but these were excluded from the totals.

Table 3

The number of users and in-game chats in the Minecraft server.

	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Number of unique players	5	18	*	24	25	27	31	50	68	71	72
Days with an in-game chat	X	X	X	X	X	7	1	5	17	8	4
Total exchanges via in-game chat	X	X	X	X	X	155	15	172	282	84	34

Note. All available data is presented here. "X" reflects data that was not collected by the server. * represents a change in server, unique users reset to zero in November. The researchers did not record the number of new users that month.

In-game chat

The number of unique players, to a limited extent, shows a willingness to play Minecraft. Attending in-person and online events is another indicator. Using English in in-game chats during unstructured play is an indication of a desire to play Minecraft in English. It is also an opportunity for an educational experience. After the researchers read the chat logs, the following types of interactions were recorded. There were 16 players who used in-game chat, most of the chat interactions involved Tetsuo, but an additional four students used the chat on at least five occasions. In February and March, chats were all S-T conversations to facilitate game cooperation. From April, interactions were S-T and S-S interactions. The S-S chats were between Japanese and Taiwanese students, welcoming them to the server and some conversations to cooperate. May had the most amount of chat interactions. There were many S-S interactions between Japanese and Taiwanese students cooperating. Other communications were basic greetings. In June, most English communication was S-T with two students and a few S-S greetings between Japanese and Taiwanese students. July saw a large decrease in interactions. These were S-T cooperation and S-S short greetings. One observation was how students often missed seeing another student by a few minutes. Being alone on the server demonstrates an inability to use English rather than an unwillingness to do so.

Other than large events, which included video calls, chat was the only known way students interacted with each other when not physically present. Video calls were possible for students through Teams and Discord, but no student-led calls

were observed. Students and teachers would also discuss Minecraft in the SLAC's conversation lounge and occasionally play together. There were four instances of Local Group Play registered on the server from April to July 2022. These involved two to three students using the SALC devices to join the Minecraft server together. Their spoken language is unknown, but the SALC encourages students to speak English. Yoshi used the option to journal after solo play twice, and the video option was never used.

Second survey

After the second event in December, the authors wished to understand why students were not joining the server initially or returning to events. Of the 24 players who had accessed the server, only five were well known to the authors and three were teachers. The second survey was given in January 2022. The researchers emailed the survey to the students from the original survey who had expressed interest in Minecraft and students who had joined the Minecraft Teams channel. This was a total of 161 students after duplicates were removed. Only 15 students responded. Eight students indicated they were too busy to join. Five students replied that they were unable to connect to the server or did not have a personal copy of the game. Only two respondents had joined, and their reason was "because it was enjoyable." The survey also revealed that survival mode was by far the preferred mode of gameplay and that there was no clear preference for communication. Nine participants, who did not play Minecraft, indicated that they did not participate in any other SALC activities. Three students indicated that they participated in 1-2 SALC activities, and three students, including the two who had joined the Minecraft server, had indicated that they participated in all activity types offered at the SALC. From these results and in light of the observations of the researchers, a website was created for onboarding future participants to the SALC Minecraft server. The SALC device count increased from one to five, and the authors decided that future events would have a group focus.

Discussion

There are several limitations to this paper. The first is that it cannot be generalized beyond participants at Sojo University because of the small participant sample size. Another limitation was the access to students. The first survey was not given to all first- and second-year students. While over half of the 243 students were interested in the server, it is unclear how many students from the pool of 1600 were uninterested and chose not to partake in the survey rather than students who were never given the survey. The second survey had a very low response rate. While the results provided ideas for how the authors might improve the server at the time, they were far from reliable, again, due to the small sample size. The questions in both surveys were trialed with other teachers, but not with students. It is possible students misinterpreted the questions. The server data on unique users only indicated the first time a user joined the server, not the total time

played or other metrics of server use; these metrics could have given a greater understanding of how students participated in the server. Finally, observations, coding, and review were all limited to the two researchers. Both researchers enjoy Minecraft and played alongside the students during this study. This presents a potential bias in the reporting of events.

The researchers witnessed collaboration and communication, both S-S and S-T. This is in line with the collaborative and communicative results of other studies (Callaghan, 2016; Egbert & Borysenko, 2019; Swier, 2014; Marcon & Faulkner, 2016; Petty, 2018; and Nebel et al., 2017). However, collaboration and communication were only observed as S-S interactions after a teacher encouraged the interaction, or after an event where communication had previously occurred between students. It appears that the Halloween and Christmas build contests did not encourage communication and were not effective activities for building a social presence. The in-person events as well as the group based Defend your Village event were a great opportunity for students to begin social interactions. These events all required a teacher to begin those social interactions before S-S interactions occurred. This is very similar to the previous studies, which had the same students participate throughout the study. The assumption that students will start a chat with another student based on a building they designed did not hold true. The teacher supported connection appears to be a great catalyst that can assist communication in unstructured play. In fact, this teacher supported focus on introductions and putting a focus on the goals of collaboration are recommended until members are comfortable in the community (Garrison, 2009). One way to improve comfort in the community could be to have an ice-breaking activity between students before beginning the main challenge of an event.

The length of time needed for these teacher-supported introductions is unknown. Events can take a long time to plan and promote. Each year there will be new students at the school, indicating that each year this support will be necessary. This is a weakness of Minecraft in the SALC, similar to the points made by McColgan (2018) and Dykes (2021); designing events can be very time consuming. While this teaching presence may continue to be necessary, the events may take less time to plan if successful events can be replicated each year. The teaching presence does not require a teacher, a student can take on that role as they mature in the community. This time commitment must also be weighed against the number of students it benefits. In the first year, no more than seven students participated in a single event. This level of participation may not warrant the time commitment required to prepare the events.

One area that was very underutilized were the written and video reflections that were given as a source of asynchronous learning. Within a classroom setting, these were found to be conducive to learning (Kuhn & Stevens, 2017; Uusi-Mäkelä, 2014; Lorence, 2015; and Smolčec et al., 2014). In this instance, the only two entries were from the most engaged student. Offering this option to participants requires very little time or effort and is in line with the SALC goal of supporting all

English skills. It may be unlikely to expect students to voluntarily use these regularly, but providing the option is worth the effort.

Conclusion

In this study, students had complete access to a SALC managed Minecraft server to use as often or as little as they wished. The Minecraft server had 72 unique players throughout the year. Many students appear to be interested in Minecraft as indicated both through surveys and general observations of students through daily interactions in class and SALC conversation sessions, but only a fraction of those players were observed using Minecraft as a means to improve their English. Planned Minecraft events saw participation as low as zero students and as high as seven students from Sojo University. A total of 16 participants used the in-game chat function. This left a high number of participants who were using the server without attending events or communicating with others in-game. Data was not available for whether these participants used the server once or many times. There are several possible reasons for this, one is that students were only interested in a single event and not the others. Another is the reduced number of in-person classes due to Covid-19 restrictions, which reduced the amount of time students spent on campus. Finally, the overlap in students interested in both English and Minecraft may be lower than believed from the first survey conducted or the motivation in some students was not high enough to overcome first-time participation anxiety. Additional research is necessary to understand which types of events will draw students into the SALC server and to better understand what can be done to lower any participation anxiety, including perceptions students hold of the English proficiency that is necessary to participate.

In terms of the educational experience, students who joined multiple events appeared to improve their English communication skills. Researchers observed longer utterances with more complex vocabulary. The researchers did not observe the desired educational experience in a few activities. The first was the building challenges, which were not conducive to building a social presence as imagined. They did not provide a natural place for communication between students or between teachers and students. Unstructured play is also not conducive to an educational experience unless a social connection has already been established between students. Group based activities were the most conducive to an education experience. These required communication to be successful. Students who join may be shy at first but have all joined voluntarily. As such, a few questions from a teacher can spark group cooperation. Further research is needed on which types of group activities are best, for example cooperative, competitive, building, fighting, or a mix of these things. Another variable to consider is whether groups who work together and are committed to several weeks of play have differing results from students who are grouped randomly with whoever is present for that event. Minecraft in the SALC appears to have value, but the best way to leverage this technology has yet to be determined.

References

- Alawajee, O., & Delafield-Butt, J. (2021). Minecraft in Education Benefits Learning and Social Engagement.. *International Journal of Game-Based Learning*, 11(4), 19–56. <https://doi.org/10.4018/IJGBL.2021100102>
- Al-Washmi, R., Bana, J., Knight, I., Benson, E., Kerr, O. A. A., Blanchfield, P., & Hopkins, G. (2014, October). Design of a Math learning game using a Minecraft Mod. In *ECGBL2014–8th European Conference on Games Based Learning: ECGBL2014* (p. 10). Academic Conferences and Publishing International.
- Baek, Y., Min, E., & Yun, S. (2020). Mining Educational Implications of Minecraft. *Computers in the Schools*, 37, 1–16. <https://doi.org/10.1080/07380569.2020.1719802>
- Bebbington, S., & Vellino, A. (2015). Can playing Minecraft improve teenagers' information literacy? *Journal of Information Literacy*, 9(2), 6–26. <https://doi.org/10.11645/9.2.2029>
- Bogost, I. (2011). *How to do things with videogames*. Minneapolis, MN: University of Minnesota Press.
- Callaghan, N. (2016). Investigating the role of Minecraft in educational learning environments. *Educational Media International*, 53(4), 244–260.
- Chien, Y. C. (2019). The Language of Massively Multiplayer Online Gamers: A Study of Vocabulary in Minecraft Gameplay. *TESL-EJ*, 23(3), 1–16. <https://tesle.org/pdf/ej91/int.pdf>
- Cipollone, M., Schifter, C. C., & Moffat, R. A. (2014). Minecraft as a creative tool: A case study. *International Journal of Game-Based Learning*, 4(2), 1–14.
- Drysdale, J. S., Graham, C. R., Spring, K. J., & Halverson, L. R. (2013). An analysis of research trends in dissertations and theses studying blended learning. *The Internet and Higher Education*, 17, 90–100. <https://doi.org/10.1016/j.iheduc.2012.11.003>
- Dykes, R. (2021). Considerations for using COTS Video Games as Language Learning Tools. *SILC Journal*, 1(1), 75–81.
- Egbert, J., & Borysenko, N. (2019). Standards, engagement, and Minecraft: Optimizing experiences in language teacher education. *Teaching and Teacher Education*, 85, 115–124. <https://doi.org/10.1016/j.tate.2019.06.015>
- Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal*, 81(3), 285–300. <https://doi.org/10.2307/329302>
- Garrison, D. R. (2009). Communities of inquiry in online learning. In *Encyclopedia of distance learning* (2nd ed.) (pp. 352–355). IGI Global.
- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3), 157–172. <https://doi.org/10.1016/j.iheduc.2007.04.001>
- Gee, J. (2013). *Good video games and good learning* (2nd ed.). Peter Lang Publishing Inc.

- Herold, B. (2015). Minecraft fueling creative ideas, analytical thinking. *Education Week*, 35(01), 12.
- Kalina, C., & Powell, K. C. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education*, 130(2), 241–250.
- Kuhn, J., & Stevens, V. (2017). Participatory culture as professional development: Preparing teachers to use Minecraft in the classroom. *TESOL Journal*, 8(4), 753–767.
- Lorence, M. (2015). School of Minecraft. *School Library Journal*. Retrieved from <https://www.slj.com/?detailStory=mincraftedu-takes-hold-in-schools>
- Marcon, N. (2013). Minecraft as a powerful literacy prompt in the secondary English classroom. *Idiom*, 49(2), 35–37.
- Marcon, N., & Faulkner, J. (2016). Exploring Minecraft as a pedagogy to motivate girls' literacy practices in the secondary English classroom. *Engineers Australia*, 51(1), 63.
- Marlatt, R. (2018). Literary analysis using Minecraft: An Asian American youth crafts her literacy identity. *Journal of Adolescent & Adult Literacy*, 62(1), 55–66.
- McColgan, M., Colesante, R., & Andrade, A. (2018). Pre-Service Teachers Learn to Teach with Serious Games. *Journal of STEM Education*, 19(2), 19–25.
- Microsoft (2021, April). *Minecraft*. https://news.xbox.com/en-us/wp-content/uploads/sites/2/2021/04/Minecraft-Franchise-Fact-Sheet_April-2021.pdf
- Nebel, S., Schneider, S., Beege, M., Kolda, F., Mackiewicz, V., & Rey, G. D. (2017). You cannot do this alone! Increasing task interdependence in cooperative educational video games to encourage collaboration. *Educational Technology Research and Development*, 65(4), 993–1014.
- Nguyen, J. (2016). Minecraft and the building blocks of creative individuality. *Configurations*, 24(4), 471–500.
- Pan Y, and Steed A. 2017. The impact of self-avatars on trust and collaboration in shared virtual environments. *PLoS ONE*, 12(12): e0189078. <https://doi.org/10.1371/journal.pone.0189078>.
- Petry, A. (2018). Playing in Minecraft: An exploratory study 1. *Revista FAMECOS*, 25(1), 1–18. <https://doi.org/10.15448/1980-3729.2018.1.27156>
- Pusey, M., & Pusey, G. (2015). Using Minecraft in the science classroom. *International Journal of Innovation in Science and Mathematics Education*, 23(3), 22–34.
- Remmerswaal, R. (2022). Minecraft in a Self-Access Language Learning Center. *JASAL Journal*, 3(1), 121–135. <https://jasalorg.com/minecraft-in-a-self-access-language-learning-center/>
- Richards, K. (2003). *Qualitative inquiry in TESOL*. Springer. <https://doi.org/10.1057/9780230505056>
- Smolčec, M., Smolčec, F., & Stevens, V. (2014). Using Minecraft for learning English. *TESLEJ*, 18(2), 1–5. <https://tesl-ej.org/pdf/ej70/int.pdf>

- Squire, K. (2011). *Video Games and Learning: Teaching and Participatory Culture in the Digital Age. Technology, Education – Connections* (the TEC Series). Teachers College Press.
- Statista (2021, October) *Number of monthly active players of Minecraft*.
<https://www.statista.com/statistics/680139/minecraft-active-players-worldwide/>
- Stevens, V. (2021). Up your game: Engage your learners with Minecraft. *TESOL Connections* (2021, September), 1–8.
http://newsmanager.commpartners.com/tesolc/downloads/features/2021/2021-09_Minecraft_Stevens.pdf
- Swier, R. (2014). Tasks for Easily Modifiable Virtual Environments. *The JALT CALL Journal*, 10(3), 203–219.
- The Community of Inquiry, (n.d.) *Framework, black & white*.
<https://www.thecommunityofinquiry.org/framework>
- Uusi-Mäkelä, M. (2014). Immersive language learning with games: Finding flow in MinecraftEdu. In J. Viteli & M. Leikomaa (Eds.), *Proceedings of EdMedia 2014--World Conference on Educational Media and Technology*. Tampere, Finland: Association for the Advancement of Computing in Education (AACE). Retrieved from
<https://www.learntechlib.org/primary/p/148409/>
- Uusi-Mäkelä, M. (2015). *Learning English in Minecraft: a case study on language competences and classroom practices* (Unpublished master's thesis). University of Tampere, Tampere, Finland.
- Vygotsky, L. (1986). *Thought and Language* (A. Kozulin, Trans.). Harvard University Press.

Authors' bios

Robert Remmerswaal is an educator in Japan with experience teaching a wide variety of ages. Currently he teaches at the university level. He is interested in ways to engage students and keep English courses interesting. His research interests include collaborative online learning, self-efficacy, and student adoption of new technology.

Robert Dykes has been living and teaching English in Japan for over a decade. He holds a master's degree in Applied Linguistics and TESOL. His main research focus is language learner motivation and anxiety, but he has presented on many topics such as CMC, grammar, and quantitative research design.

4

Ready, set...? End-user testing of an AR treasure hunt learning game

Samuel Taylor and Adam Stone, Kyushu Sangyo University

Abstract

Formative user testing, aimed at obtaining feedback from students, instructors, and administrators at a higher education institution in Japan, was undertaken on an AR treasure hunt learning game app. It involved the collection of usage data from data-mining tools, anecdotal evidence from surveys and interviews, and descriptive evidence from observations of tester behavior to investigate aspects of app learning, engagement, usability. Analysis indicated that testers ($N = 21$) rated learning significantly higher than both engagement ($p < 0.01$) and usability ($p < 0.01$), and that engagement was rated significantly higher than usability ($p < 0.05$). Observations noted indicators of engagement, frustration, and unanticipated patterns of usage. Interviews with testers raised issues concerning usability and engagement, and provided suggested changes to the app. These were used to make improvements that, while impacting learning, also facilitated the classroom implementation of the app in pedagogy. In conclusion, it is suggested that user testing is a vital part of educational technology impact evaluations, and that key aims of the design and development of educational apps should be the easy access, enabled through high usability and engagement, of learning content situated in contexts that provide opportunities for further or applied use.

本研究では、日本の高等教育機関の学生、講師、管理者からのフィードバックを得ることを目的とした形成的ユーザーテストを、AR宝探し学習ゲームアプリで実施した。データマイニングツールによる利用データの収集、アンケートやインタビューによる逸話的証拠の収集、調査対象者の行動観察による記述的証拠の収集により、アプリの学習、エンゲージメント、ユーザビリティの側面を調査した。分析によると、調査対象者 ($N=21$) は学習をエンゲージメント ($p < 0.01$) およびユーザビリティ ($p < 0.01$) の両方よりも有意に高く評価し、エンゲージメントはユーザビリティ ($p < 0.05$) よりも有意に高く評価されたことが分かった。また、観察では、エンゲージメント、フラストレーション、そして予期せぬ使用パターンが指摘された。調査対象者へのインタビューでは、ユーザビリティとエンゲージメントに関する問題点が指摘され、アプリの

変更案が提示された。これらは、学習に影響を与える一方で、教育上、アプリを教室で容易に実施するために使用された。結論としては、教育工学におけるインパクト評価には、ユーザーテストが不可欠であり、教育用アプリの設計・開発の主な目的は、高いユーザビリティとエンゲージメントによって可能になり、さらなる活用や応用の機会を提供する上で、学習コンテンツへの容易なアクセスであることを提案した。

Keywords: Mobile learning, Augmented Reality, Digital game-based learning, User testing, Educational apps

Introduction

“KSU Treasure Hunter” is a campus guide smartphone application (app) designed and developed by the authors at Kyushu Sangyo University (KSU), Fukuoka, Japan. The app is part of a research project (see Taylor, 2020; Taylor & Stone, 2018; Taylor et al., 2019) investigating how Augmented Reality (AR) technology can facilitate active participation in Japanese higher education to meet the goals of acquiring lifelong learning abilities and independent thinking skills (Japanese Ministry of Education, Culture, Sports, Science, and Technology, 2012). The project assumes that educational technology is not simply a tool for sustaining or streamlining prevailing methods of instruction (Christensen, 2008) but a means of transforming education by supporting new kinds of relationships between learners and technology (An & Oliver, 2020). In this context, AR is considered an emerging technology as it contains the potential to change the experience of education (Cukurova & Luckin, 2018). The creation and use of AR materials is also informed by a constructivist view of learning, in which learners construct their understanding through their experiences so that learning is active and exploratory, authentic and contextualized (Pegrum, 2014).

KSU Treasure Hunter presents practical information about campus services in a bilingual, pirate treasure hunt-themed AR learning game. Posters placed in campus facilities act as AR markers, requiring students to visit facilities to learn about available services. As such, the KSU Treasure Hunter app aims to transform the experience of learning about campus for KSU students. Its design is less concerned with improving the efficiency of knowledge transmission than with enabling learners to engage with the world in new ways. The magic of AR in revealing some kind of previously hidden truth (Norman, 2018) in exploratory and contextualized learning experiences increases learner interest (Godwin-Jones, 2016), and contributes to a positive learner experience (Taylor & Stone, 2018). The AR reveals campus information in the form of a learning game. Instructional goals are mapped onto the first three levels of Bloom’s taxonomy (Armstrong, 2019) in a 3-level structure, detailed in Taylor (2020), that engages app users in progressively higher-level thinking skills. As such, pedagogy is embedded in a fun learning process that involves an abstraction of reality or an element of fantasy in the teaching process (Boller & Kapp, 2017). The learning game players learn the names of campus facilities and services, understand what the services entail, and apply knowledge of the services in order to solve problems commonly encountered by students. Table 1 provides an example of the learning goals, learner

actions, and examples of learning content for one service available at the KSU library.

Table 1

The learning game goals, learner actions, and examples of learning content for one service available at the KSU library

		Level 1	Level 2	Level 3
Learning goal		learn the names of campus facilities and services	understand what the service entails	apply knowledge of the services in order to solve problems commonly encountered by students
Learner action		match the name of a service presented by the pirate NPC with the name of the same service presented on a jewel rotating around the AR marker	match a description of a service presented by the pirate NPC with the name of the same service presented on a jewel rotating around the AR marker	match a problem presented by a student NPC with the name of the same service presented on a jewel rotating around the AR marker
Learning content	Prompt	career guidance literature	Preparation is everything... make informed decisions on your career choices, and know what you need to do to get the job you want.	I would like some information on jobs; I want to know if I need any special qualifications to get the kind of job I'm interested in.
	Correct response		career guidance literature	

The testing of emerging educational technology is part of iterative cycles of testing, design, development, and evaluation aimed at achieving incremental improvements to learning and teaching (Cukurova & Luckin, 2018). User testing, 'a group of usability evaluation methods that involve user participation' (Riihiaho, 2018, p. 257), is recognized as a critical part of the development process (Crowther et al., 2004). For educational apps, user testing should include evaluations of usability, engagement, and learning (McQuiggan et al., 2015). Testing involves the collection of anecdotal evidence gained from surveys and interviews with testers, and descriptive evidence gained from observational note-taking of user behavior (Cukurova & Luckin, 2018) and indicators of enjoyment, engagement, and frustration (Sim et al., 2006).

It is traditional to assess usability by taking measures of the users' performance, by noting where users have difficulties with the interface, and by asking the users for their opinions of the product (Sim et al., 2006). A small number of testers, between five and ten, are sufficient for identifying problems (Riihiaho, 2018). Tests should be realistic and represent actual use (Riihiaho, 2018), however testing in lab conditions as opposed to in the field has several advantages: it is less time-consuming and leads to fewer interruptions and, most importantly, there is no difference in the number of user interface problems identified by testers (Kaikkonen et al., 2005). Testing can also involve analysis of in-game user interactions, obtained from data mining tools built into the app, to investigate whether users are showing expected sequences of behavior and whether there are any over- or underused features (McQuiggan et al., 2015).

Ferrer et al. (2013) user tested an AR game to investigate usability, learning, and motivation. Usability concerned the design of the AR, and analysis indicated that, despite the AR game having reduced usability compared to a desktop version because it took longer to play, the AR game was more motivating for learners. Learning was found to be unaffected by mode. The measurement of the motivation of the college-student testers was achieved using survey responses. Learning was measured using pre-post measures. Kucirkova et al. (2014) evaluated engagement with educational software using analysis of user interactions with the software and with each other. It was demonstrated how the engagement of nursery school-aged children with a Spanish language learning app was shown to be affected by both usability and learning content. In particular, it was suggested that ease of access to learning content that encourages open-ended accomplishments, rather than drill-and-practice activities, facilitates greater engagement.

Method

The testing of KSU Treasure Hunter was largely concerned with usability, particularly the user interface (UI), as such, it can be described as formative, as it was aimed at obtaining feedback from users to inform further development (Riihiaho, 2018). The testing detailed in this paper was the last round of testing before the app was made available for use by instructors and students. It involved the collection of usage data from data-mining tools in the app, anecdotal evidence from surveys and interviews, and descriptive evidence from observations of tester behavior.

Testers from three distinct stakeholder groups were recruited: students, instructors, and administrators. This was done to obtain feedback from the three populations involved in the use of the app upon it being made available, each of which may have different insights on the app's objectives and limitations (McQuiggan et al., 2015). A total of ten Japanese students participated voluntarily in testing, and their participation was not connected to their English classes or English proficiency in any way. The instructor group consisted of five English instructors – two Japanese, two English, and one American – and the administrator group consisted

of six Japanese administrative university employees. Demographic data, such as age and gender, were not recorded, and testers were assigned a letter of the alphabet, from A to U, to allow for anonymity.

Testing was conducted in lab conditions, in a single room on the university campus. The placing of the AR markers in the room meant that testers could experience the app content without having to visit campus facilities. Participants tested the app on their own devices and were free to choose in which language, English or Japanese, they used the app. As they concern usability, the operating system (OS) of the device used by each tester, and whether the tester used the app in English or Japanese and with or without sound, were recorded. Information on the first two items was retrieved from in-app data-mining tools, whereas the use of sound or otherwise was noted down on paper by a researcher during observations of tester behavior. Guidance was only provided when it concerned accessing the app by downloading it onto the tester’s device. There was no time limit enforced for testing, and testing was considered to be complete when each tester perceived that the app experience had finished. Testers were then given the survey to complete. After that, interviews were conducted in either Japanese or English by a researcher, depending on the tester’s preference, and were audio recorded. The researcher conducting the interviews also took notes of tester responses. Tester suggestions for improvements made during the interviews were listed and categorized by whether they aligned with aspects of app learning, usability, or engagement, and coded by aspects of the app design and development, such as UI, AR design, system architecture, game design, and feature design.

Following McQuiggan et al. (2015) and Sim et al. (2006), the survey aimed at measuring tester opinions on app learning, usability, and engagement, and their overall opinion of it. Consequently, there were three, 5-point Likert items, shown in Table 2 below, for learning, usability, and engagement, and one, 5-point app rating item. The learning item asked to what degree the app is good for learning, the usability item asked to what degree the app is easy to use, and the engagement item asked to what degree the app is fun to use.

Table 2

Survey used in testing

Item	Likert scale				
This app is good for learning about campus.	1	2	3	4	5
This app is easy to use.	(strongly disagree)	(disagree)	(neither)	(agree)	(strongly agree)
This app is fun to use.					
Please rate this app out of five.	1	2	3	4	5
	(very bad)	(bad)	(neither)	(good)	(very good)

The survey data were analyzed, using independent samples *t* tests, for significant differences between the scores of the four survey items for the whole group of testers (N=21), across the three tester groups, and across three distinct usage groups of OS, language, and sound. This was done to investigate whether the app was significantly stronger or weaker in perceptions of learning, usability, and engagement, whether the tester groups experienced the app significantly differently, and whether different usage patterns significantly influenced perceptions of the tester experience.

The 19 interview questions were informed by the app checklist in McQuiggan et al. (2015), and decided upon by the researchers after consideration of the primary concerns regarding the continued design and development of the app. Table 3 below displays the questions split into categories of app purpose, learning, usability, and engagement. The questions were shared with testers without being split into categories. The Japanese translations of the questions are provided in Appendix A. Tester responses to the interview questions were reviewed to identify any specific suggestions made regarding changes to the app. These suggestions were considered in tandem with the observations of tester behavior to create a list of actionable changes.

Table 3

The interview questions used in testing

Purpose	
1	Do you think the app works well as a campus guide?
Learning	
2	Do you think that game progress is too easy or too difficult?
3	Do you think that game progress is too slow or too quick?
4	Do you think that the game provides enough feedback?
5	Do you think that the game encourages the player to think about how to use campus services?
6	Do you think that the game experience is relevant to real life?
7	Do you think that the game is an efficient way of learning about campus services?
8	Do you think that the game is an engaging way of learning about campus services?
Usability	
9	Were there any moments when the app was difficult to use?
10	Were there any moments when navigating through the app was unintuitive?
11	Were there any moments when the text was difficult to read?
12	Do you think that the app offers enough support on how to use it?

Engagement

- 13 What parts of the app were the most fun?
 - 14 What parts of the app were the least fun?
-

General

- 15 What is a good point of this app?
 - 16 What is a bad point of this app?
 - 17 Is there anything about the app that you think should be changed?
 - 18 Would you use this app again?
 - 19 Do you have any other opinions on the app?
-

Results

Survey data

Survey data, shown in Table 4 below, showed that testers (N=21) rated learning the highest (5.0) and usability the lowest (4.0). Learning was rated significantly higher than both engagement ($p<0.01$) and usability ($p<0.01$), and engagement was rated significantly higher than usability ($p<0.05$). In the separate tester groups, administrators (n=6) rated the app significantly higher than students (n=10) ($p<0.01$) and instructors (n=5) ($p<0.05$), and rated engagement significantly higher than students ($p<0.05$) and instructors ($p<0.05$). There were no other significant differences.

Table 4

Ratings of app learning, engagement, and usability for total (N=21), student (n=10), instructor (n=5), and administrator (n=6) tester groups

	Total (N=21)	Student (n=10)	Instructor (n=5)	Administrator (n=6)
Learning	5.0	4.9	5.0	5.0
Engagement	4.5	4.4	4.0	5.0
Usability	4.0	3.8	4.0	4.0
App rating	4.5	4.3	4.2	5.0

Usage data

The majority of testers tested the app in Japanese (n=15, 71%), using iOS (n=14, 67%) with the sound off (n=13, 62%). Testers who initially used the app with no sound and then turned on the sound mid-testing are recorded under 'Both'. The

student tester group largely reflected the whole group (Japanese: n=9, 90%; iOS: n=9, 90%; sound off: n=8, 80%), whereas the instructor and administrator groups had a higher proportion of Android OS (n=3, 60%; n=3, 50%) and sound usage (n=3, 60%; n=3, 50%). The instructor group had the highest proportion of English usage (n=4, 80%). The data is provided in Table 5, below. Despite these variations there were no significant differences between language, OS, and sound usage groups in survey scores of learning, usability, engagement, or overall app rating.

Table 5
Tester usage data

Group	Language		OS		Sound		
	English	Japanese	iOS	Android	Yes	No	Both
Total (N=21)	6 (29%)	15 (71%)	14 (67%)	7 (33%)	6 (28%)	13 (62%)	2 (10%)
Student (n=10)	1 (10%)	9 (90%)	9 (90%)	1 (10%)	0 (0%)	8 (80%)	2 (20%)
Instructor (n=5)	4 (80%)	1 (20%)	2 (40%)	3 (60%)	3 (60%)	2 (40%)	0 (0%)
Administrator (n=6)	1 (17%)	5 (83%)	3 (50%)	3 (50%)	3 (50%)	3 (50%)	0 (0%)

Observations

Tester behavior indicated signs of both engagement and frustration with the app. Engagement, indicated by comments and laughter, was most often observed when testers used the game introduction scene, in which the player first meets the AR NPC guide character. Frustration, indicated by sighs and looking around the room, was observed when testers were in the three-level learning game. Two testers exhibited unanticipated patterns of usage by not using a button prompt on a level summary screen to continue to level two after completing level one. To continue with the game these testers quit the app and reopened it, continued their game from the main menu, accessed the next location from the map menu, and completed only the first level for each location.

Interviews

Testers were asked about indicators of frustration and it was reported that the learning game was too repetitive, and that level one, in particular, could be made more interesting by changing the game mechanics. Other issues raised concerned usability and engagement. Table 6, below, displays tester-made suggested changes, categorized by the user experience aspects of usability and engagement. There were nine suggestions relating to usability, five relating to engagement, and none relating to learning. Inferred by the researchers, the suggestions are

Table 6

Suggested changes provided by testers and the underlying problems inferred by the researchers

Suggested change	Underlying problem	Aspect of design and development
Usability		
Use a real campus map.	User can't use map for function of locating campus services.	Feature design
Make the campus information screen more prominent.	User not using app as campus information guide.	Feature design
Make the campus information screen look different to the map screen.	The function of the campus information is unclear.	Feature design
Put menu button labels in Japanese in the Japanese version.	User can't understand what each menu button text means.	UI
Change English font to one easier to read.	Menu and learning game content is difficult to read.	UI
Add AR use guidance.	AR content can be difficult to access using AR marker.	AR design
Make some game content playable in device.	Holding up phone to play AR game is tiring. User is tied to AR marker.	AR design
Add a menu button to the language selection screen.	User unable to view menu before beginning a game.	System architecture
Add a menu button in AR game.	User unable to view menu while in AR game.	System architecture
Engagement		
Add greater reward for completion of learning game.	User unengaged by experiential design.	Game design
Change color of AR game feedback text to indicate if response is correct or incorrect.	AR game learning feedback is unclear.	Game design
Indicate completion of a location more clearly on the map screen.	AR game progress feedback is unclear.	Game design
Add variety to AR game.	AR game mechanics become repetitive.	Game / AR design
The pirate NPC should speak	Pirate NPC not as prominent as it could be.	AR design

accompanied by description of the underlying problem that prompted the suggestion, and the aspect of the app design and development that each suggestion concerns. The suggestions relating to usability concerned feature design, UI, AR design, and system architecture. In contrast, the suggestions relating to engagement largely concerned game and AR design, including game learning and progress feedback and game mechanics.

Discussion

Learning

In general terms, analysis of survey and usage data indicated that instructors and students had similar experiences of the app, while administrators experienced the app more positively than the other two groups, seen in higher scores for app engagement and overall rating. However, learning was the most positively experienced aspect of the app, and was not experienced differently by tester group or usage pattern. These findings suggest that the learning content is sufficient for the intended purpose of the app. Alternatively, it may also indicate that formative user testing is not the best means of collecting evidence on the impact of an emerging technology on learning.

Usability

Usability issues were shown to exist by usability being the least positively experienced aspect of the app. The usability issues were perceived by all tester groups, but they weren't distinguished by device OS, language, or sound usage. These results can be seen to reflect the tendency for AR content to foster a positive learner experience (Taylor & Stone, 2018) and motivate students to learn (Ferrer et al., 2013) despite the presence of usability issues. Analysis of the suggested changes gained from the interviews with testers supports the survey and observational data in indicating that usability issues outnumber those concerning engagement or learning. More suggestions were made that relate to usability than engagement. The most frequently occurring aspect of the app design and development in suggestions related to usability was feature design. The underlying problem of each feature design issue concerned expanding the scope of the app from a learning game to a campus guide. This suggests that the usability of the learning game is affected by the extent to which it is integrated into part of a larger campus information system. The provision of open-ended accomplishments that encourage high degrees of engagement and increase the educational value of the app (Kucirkova et al., 2014) needs to explicitly concern this context. Open-ended accomplishments for a campus knowledge system can be encouraged using a mobile app's ability to connect the user to the wider world. The challenge is to integrate activities, such as booking an appointment for a service, using links to external information related to campus services, such as resume writing tips, or using links to KSU sites, such as the LERC Moodle system, into the learning game

so that learning is active and exploratory, authentic and contextualized (Pegrum, 2014).

As a result of formative user testing, changes made to the app to improve usability concerned feature design and UI. The campus information screen was redesigned to be more distinct from the campus map screen to clarify the apps intended use as a campus guide. UI buttons were redesigned to be more clearly recognizable as interactive features to avoid the incorrect patterns of usage described above. The English font was changed to improve the clarity of learning content. All app content was made bilingual to improve the accessibility of learning content. Figure 1 below displays the changes to the UI buttons and English font on the game level three summary screen.

Figure 1

The changes to the UI buttons and English font as a result of user testing



Engagement

Engagement was the most positively experienced aspect of the app after learning, and it was more positively experienced by administrators than by the other tester groups. It was not affected by sound usage, device OS, or language use. The most frequently occurring aspect of the app design and development in the suggestions relating to engagement was learning game design. Therefore,

learner actions, determined by game mechanics, and learning feedback can be seen as particularly important for user engagement in learning games, as they make learning content accessible, engaging and comprehensible. The issue of the accessibility of learning content, the importance of which is demonstrated by Kucirkova et al. (2014), also concerns usability, and the results of testing suggest that accessibility is clearly related to app feature design, UI, AR design, system architecture, and game design.

The suggested changes regarding learning game feedback were enacted, so that green and red text was used in learning feedback to indicate correct and incorrect responses. In regard to reducing the repetitiveness of the learning game, the following modification to level one was proposed but not enacted due to time constraints. The AR marker gives access to a grid of 15 jewels, which are not tied to the AR marker but are playable in-device. The pirate NPC presents the task facing the player: to find and select three of each service as quickly as they can. Each set of three jewels is collected upon being selected in a sequence. This correct response also receives a point score, but incorrect responses, such as selecting a different service to the one selected previously, lead to a points deduction. A timer is added above the grid of jewels, and the time taken for the player to complete the level is used with the number of correct and incorrect responses to calculate a final score. This proposed change is shown in Figure 2 below. Adding time pressure is a means of increasing engagement, whereas the use of game content not tied to the AR marker improves usability by not requiring the player to hold their device in front of the AR marker for the duration of the game level.

Figure 2

A possible modification to learning game level one to improve usability and engagement

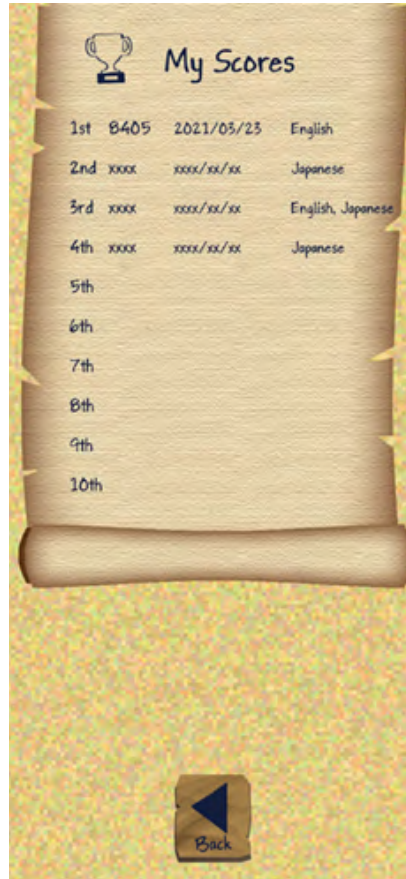


Pedagogy

The changes detailed above all modify the app design, and in-app experience and behaviors, and, therefore, can be seen to concern learning (Cukurova & Luckin, 2018). There was also a change made to the app that concerned pedagogy. A 'My Scores' screen, shown in Figure 3 below, was added to the high scores that displays only the device scores. While this adds app personalization, an important feature of educational apps (McQuiggan et al., 2015), this screen also displays in which language the app was played. This information is included as a means for instructors to check that learners have used the app appropriately when implemented in classroom use.

Figure 3

The newly-added My Scores screen



Conclusion

The current study provides anecdotal and descriptive evidence of the development and testing stage of the KSU Treasure Hunter app. There are several drawbacks of the current research that temper any implications that can be made regarding the results of analysis. Firstly, it is necessarily specific to the context of KSU in learning content, and by not adopting any of the various frameworks available for the evaluation of educational apps (see McQuiggan et al., 2015) the research does not allow for the comparison of KSU Treasure Hunter with other similar apps. Secondly, a simple testing design makes contributing to usability testing (see Riihiaho, 2018), and the measurement of user engagement (see Kucirkova et al., 2014) and learning (see Ferrer et al., 2013) difficult.

Despite these limitations, the research has shown that user testing is a vital part in achieving improvements to the learning and teaching associated with emerging educational technology (Cukurova & Luckin, 2018). Following Riihiaho (2018), testing in lab conditions with a small number of testers was sufficient for

identifying problems with the app in regard to both usability and engagement. As suggested by Sim et al. (2006), asking user opinions of the app was also a productive means of informing design and development. The combination of survey, observational, and interview data provided the app designers and developers with a roadmap for future iterations, but also suggested that issues concerning learning may be best explored through other types of data collection. Future research concerned with learning could involve the collection of correlational data, for example on the relationship between user perceptions of the app and user performance as indicated by game scores. Additionally, the collection of causal evidence on the pedagogical soundness of the learning game could be gained from the classroom implementation of the app in a pedagogical sequence using pre and post-test data. In this way, future research can take advantage of the opportunity the app provides to engage in a holistic approach to an emerging technology impact evaluation that includes various types of evidence of the effect of the app on learning and pedagogy (Cukurova & Luckin, 2018).

KSU Treasure Hunter is a unique instructional and promotional tool by its integration of a campus knowledge system and an AR learning game. Formative testing suggests that key aims of the design and development of educational apps should be the easy access of learning content, that access is enabled through high usability and engagement, and that learning content should be situated in contexts that provide opportunities for further or applied use. From this perspective, the impact of KSU Treasure Hunter can be improved by enhancing campus knowledge system features and improving the learning game to include more variety in gameplay and the possibility of contextually situated open-ended accomplishments. Whether these adjustments bring the app closer to enabling a transformation of education (An & Oliver, 2020) is, as yet, unanswered.

References

- An, T. & Oliver, M. (2020). What in the world is educational technology? Rethinking the field from the perspective of the philosophy of technology. *Learning, Media and Technology*, 46(1), 6–19.
<https://doi.org/10.1080/17439884.2020.1810066>
- Boller, S. & Kapp, K. (2017). *Play to learn: Everything you need to know about designing effective learning games*. ATD Press.
- Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disruptive innovation will change the way the world learns*. McGraw-Hill.
- Crowther, M. S., Keller, C. C., & Waddoups, G. L. (2004). Improving the quality and effectiveness of computer-mediated instruction through usability evaluations. *British Journal of Educational Technology*, 35(3), 289–303.
- Cukurova, M. & Luckin, R. (2018). Measuring the impact of emerging technologies in education: A pragmatic approach. In J. Voogt et al (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 1181–1199). Springer International Publishing.
https://doi.org/10.1007/978-3-319-71054-9_81

- Ferrer, V., Perdomo, A., Rashed-Ali, H., Fies, C., & Quarles, J. (2013). How does usability impact motivation in augmented reality serious games for education? *2013 5th international conference on games and virtual worlds for serious applications (VS-GAMES)* (pp. 1–8).
<http://dx.doi.org/10.1109/VS-GAMES.2013.6624233>
- Godwin-Jones, R. (2016). Augmented reality and language learning: From annotated vocabulary to place-based mobile games. *Language Learning and Technology*, 20(3), 9–19.
- Japanese Ministry of Education, Culture, Sports, Science, and Technology. (2012). *Towards a qualitative transformation of university education for building a new future - Universities fostering lifelong learning and the ability to think independently and proactively*. Retrieved October 22, 2018, from:
<http://www.mext.go.jp/en/publication/report/title01/detail01/1380275.htm>
- Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2005). Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability Studies*, 1(1), 4–16.
- Kucirkova, N., Messer, D., Sheeny, K., & Panadero, C. F. (2014). Children’s engagement with educational iPad apps: Insights from a Spanish classroom. *Computers and Education*, 71, 175–184.
<https://doi.org/10.1016/j.compedu.2013.10.003>
- McQuiggan, S., Kosturko, L., McQuiggan, J., & Sabourin, J. (2015). *Mobile learning: A handbook for developers, educators, and learners*. John Wiley & Sons, Inc.
- Norman, J. (2018). L. Frank Baum’s “The Master Key” imagines a kind of augmented reality.
<http://www.historyofinformation.com/expanded.php?id=4698>. Accessed November 14th, 2018.
- Pegrum, M. (2014). *Mobile learning: Languages, literacies and cultures*. Palgrave Macmillan.
- Riihiaho, S. (2018). Usability testing. In K. L. Norman & J. Kirakowski (Eds.), *The Wiley Handbook of Human Interaction* (Vol. 1). John Wiley & Sons Ltd.
- Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers and Education*, 46, 235–248.
- Taylor, S. (2020). Designing an augmented reality-enabled smartphone campus guide learning game. In P. Fotaris (Eds.) *Proceedings of the 13th international conference on game based learning ECGBL 2020* (pp. 604–612).
- Taylor, S. & Stone, A. (2018). Smartphone augmented reality for EFL learners. *2018 PanSIG Journal*, 258–263.
- Taylor, S., Stone, A. & Witkin, N. (2019). Developing an educational and promotional augmented reality learning game smartphone application. *EdMedia + Innovative Learning*, 2019(1), 935–938.

Appendix

The Japanese version of the interview questions used in testing.

1. アプリはキャンパスガイドとして適していると思いますか？
2. ゲームの進行について簡単すぎると思いますか、それとも難しすぎると思いますか？
3. ゲームの進行について遅すぎますか、早すぎますか？
4. ゲームは十分なフィードバックを提供していると思いますか？
5. ゲームは学内施設利用情報について考えるきっかけになると思いますか？
6. ゲームの体験は実生活に関連していると思いますか？
7. このゲームでは学内施設利用情報について効率的に学ぶことができると思いますか？
8. このゲームは学内施設利用情報について学ぶための魅力的な方法だと思いますか？
9. アプリが使いづらかった場面はありましたか？
10. アプリの操作が直感的でない場面はありましたか？
11. テキストが読みにくい場面はありましたか？
12. 使い方のサポートは十分だと思いますか？
13. アプリのどの部分が一番面白かったですか？
14. アプリのどの部分が一番面白くなかったですか？
15. アプリの良い点は何だと思いますか？
16. アプリの悪い点は何だと思いますか？
17. アプリについて、変更すべきだと思う点がありますか？
18. 再度このアプリを使いたいと思いますか？
19. アプリについて、他に何かご意見はありますか？

Authors' bios

Samuel Taylor received his MA in TESOL with distinction from the Institute of Education, University College London in 2017. His research interests include speaking, student perceptions of education, and educational technology.

Adam Stone received his Master of Applied Linguistics from Massey University, New Zealand in 2016. His research interests are English vocabulary acquisition, Computer Assisted Language Learning (CALL), Mobile Learning, educational software design and development, and Mixed, Augmented and Virtual Realities (MAVR).

5

Evaluation of a VR language learning environment: Effect of feedback on learners' flow state

James York, Meiji University
Hirotaka Takuma, Takuya Yoshino and Hiroshi Nakayama, Tokyo
Denki University

Abstract

Language learning motivation may be fostered by inducing a “flow state” in learners. This is characterised by a state of deep immersion in an activity, such as feeling enjoyment and satisfaction in the activity itself. In this study, the potential of virtual reality (VR) to produce learners' flow state is the research focus. Investigations relate to whether adding audiovisual feedback to a pairwork speaking activity can promote flow. A pairwork spot-the-difference activity that utilized the playful and interactive affordances of VR was created. Two versions of the environment were created: one with audiovisual feedback, and the other without. 22 participants, separated into two groups experienced the VR environment with ($n = 12$) and without feedback ($n = 10$). A questionnaire with 10 measures was used to determine whether the VR environments facilitated flow (based on Cho, 2018). Results of the questionnaire suggested that there was no significant difference in the flow state of the participants with and without feedback. However, examination of individual measures revealed significant differences in mean scores for two measures: both “enjoyment” and “satisfaction” were higher in the group that experienced the VR environment with feedback, suggesting that feedback in VR may promote motivation. However, due to the low number of participants in this study, the generalization of results is difficult.

言語学習の動機付けは、学習者の「フロー状態」を誘発することで促進される可能性がある。フロー状態とは、ある活動に深く没頭し、その活動自体に楽しさや満足感を感じる状態のことである。そこで本研究では、学習者のフロー状態を作り出すバーチャルリアリティ (VR) の可能性に着目し、ペアワークでのスピーキング活動に視聴覚フィードバックを加えることで、フローを促進できるかどうかを調査した。本研究ではVRの遊び心とインタラクティブ性を活用したペアワークでの間違い探しシステムを開発した。

本実験では、視聴覚フィードバックがあるものと無いものの2種類のVR環境を開発した。22名の参加者

は2つのグループに分かれ、フィードバックあり(n = 12)とフィードバック無し(n = 10)のVR環境を体験した。VR環境がフローを促進したかどうかを判断するために、10の尺度で構成されたアンケート(Cho , 2018)を使用した。

アンケートの結果、フィードバックの有無による参加者のフロー状態に有意な差はないことが示唆された。しかし、尺度を個別に検討した結果、2つの尺度の平均値に有意差があり、「楽しさ」「満足感」ともにフィードバックありのVR環境を体験したグループで高く、VRにおけるフィードバックがモチベーションを促進する可能性が示唆された。

本研究は参加人数が少ないため、結果の一般化は困難である。

Keywords: virtual reality, flow state, game-based learning, motivation

Introduction

English, as a global lingua franca is being used increasingly within Japan. Additionally, the recent hosting of the Olympic and Paralympic Games in Tokyo has influenced the importance of improving English proficiency for Japanese nationals. Indeed, the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) has pushed for the further development of Japanese students' communication skills (MEXT, 2014). However, many Japanese students still have a poor command of English and struggle to find the motivation to learn.

Although there are various methods of measuring learner motivation (see Dornyei & Ushida, 2011), this study focuses on the development of a *flow state* using the affordances of a specific technology (Csikszentmihalyi, 1990). The flow state is a condition of deep immersion in an activity (Nakamura & Csikszentmihalyi, 2014). It results in a high level of concentration and a feeling of enjoyment and satisfaction. It is also characterized by a distortion of the sense of time such that the time experienced by an individual feels shorter than the actual elapsed time. There are three theorized ways in which flow may be promoted (Csikszentmihalyi, 1996):

- 1 **Balancing** the level of **challenge** with a participant's skill level,
- 2 the provision of **immediate feedback** on an activity performed by the participant,
- 3 and making sure the participant is aware of the goal they are working towards. Hence, the provision of **clear goals**.

Flow in learning and education research

Flow is considered an important construct within educational contexts for the following reasons. First, flow may promote increased task engagement as students exert attention on learning activities. This is hypothesized to lead to better learning outcomes (Christenson et al., 2012). Secondly, flow, which occurs when learners are provided with a balanced level of challenge in accordance with their skill level, may be considered the optimal condition for effective, deep learning (Hamari et al., 2016). Thirdly, as flow provides learners with a positive experience, it is also hypothesized that those who have experienced flow will want

to experience it again and will therefore set higher challenges for themselves (keeping the skill-challenge balance optimal), and thus work harder (Engeser & Rheinberg, 2008). In summary, flow theory claims that if the underlying triggers for flow are provided, people push themselves to “higher levels of performance” (Csikszentmihalyi, 1990, p. 74).

Since the flow state is an intrinsic motivator and a factor that promotes the growth and development of abilities, examinations of how flow can be generated in students as they engage with educational content have been explored in various research fields. In relation to the current study, we are particularly interested in how interactive environments or “games” may be developed to promote flow. Perttula, et al., (2017) conducted a systematic literature review of “game-based learning” (henceforth: GBL) studies. One of the major findings of their review was that flow had a positive effect on enhancing players’ performance, overall learning gains, and engagement with environments. Additionally, of the 19 studies they examined, “there was not a consistent way to be found to measure the flow experience among the papers” (p. 62). In general, however, flow was measured via a questionnaire given to participants after playing a game.

Language learning and flow

In keeping with the goal of the current study, there are also a number of papers that have explored the generation of flow in foreign language learning contexts.

Egbert (2003) compared the experience of flow to that of Krashen’s (1982, p. 66) Forgetting Principle which states that “the best input is so interesting and relevant that the acquirer may even ‘forget’ that the message is encoded in a foreign language.” This early paper on flow in SLA also mentions the use of games and simulations as a promoter of flow due to the user-controlled pace and content as well as its immersive nature which hooks learners into the learning world. Egbert also created a highly detailed model of how flow may be generated through the design of language tasks including the following (p. 502):

- Appropriate challenge
- Making the task interesting for learners
- Providing enough time for completion
- Providing immediate feedback
- Making sure learners feel that they are in control
- Giving learners the opportunity to focus without interruption

Note that many of these elements overlap with factors in Robinson’s (2007) Triadic Componential Framework such as planning time (task complexity factor) and task motivation and openness to completing the task (learner factors). Thus, flow and task design are intimately entwined in that designing tasks to meet certain learner characteristics are the same foundations for generating flow.

Continuing with Egbert’s study, flow was measured in terms of a learner’s *control* over a task, *attention* paid towards the task, and *interest* in completing the task itself where results suggest that certain tasks promoted flow more than others;

yet Egbert was left with the observation that “it is also clear that we cannot fully explain [flow generation]” (p. 513) in that it was difficult to ascertain which task elements promoted flow.

Subsequently, Cho (2018) used four argumentative tasks of different task complexity and modality (spoken and written modes) to measure how task design affected flow. The level of challenge for each task was operationalized in terms of the number of elements that students had to manipulate (+/- number of elements, a *task difficulty* factor in Robinson’s Triadic Componential Framework) as well as by the modality of the tasks. Flow was measured based on the same constructs as Egbert: control, attention, and interest in a task. 141 learners conducted the four tasks where it was found that rather than the task design construct, only modality had a statistically significant effect on whether participants experienced flow. Specifically, the writing tasks elicited a sense of challenge-skill balance which was found to be the most significant predictor of flow in this study. Thus, and in conclusion, the speaking tasks in Cho may have been more effective in generating flow if their task difficulty was reduced in comparison to the written tasks as the written mode is less cognitively demanding and may therefore allow learners to commit more attentional resources towards task goals (Kormos, 2014; York, 2019).

Virtual reality and flow

The present study is concerned with the effect of virtual reality (VR) technology on flow in language learning. Previous studies have examined how VR may alleviate foreign language anxiety (York et al., 2021), as well as its effect on learners’ oral output (Tokutake et al., 2021) and motivation towards studying English (Shibata & York, 2021). In the present study, an environment very similar to that of York et al., (2021) was created. Results of that study suggested that compared to other modes of communication (audio-only, and video-based chat), participants found the VR domain to be 1) the easiest mode within which to communicate with an interlocutor, 2) the most enjoyable, and 3) the most effective of three domains for language learning.

Regarding flow in particular, one study found that VR amplifies the flow state (Kim & Ko, 2019), and another showed that compared to playing a game on a 2D screen, in the VR domain time passed faster for participants (Rutrecht, et al., 2021). This aligns well with previous literature on flow where, as mentioned in the introduction, individuals perceive elapsed time to be shorter than the elapsed time when experiencing a flow state. Additionally, in relation to York, et al., (2021), the results of Rutrecht et al. (2021) suggest that the immersive and fun, game-like nature of VR may be more effective than other modes in promoting a flow state.

Research questions

Based on a review of the literature on flow, VR, and language learning contexts, the following research questions were formulated:

- 1 Does the provision of feedback in a language learning VR game promote a flow state?
- 2 Do participants' perceptions of elapsed time and actual elapsed time differ between the two environments?
- 3 What are participants' perceptions of learning English with a VR environment?

Method

The purpose of this study is to compare the state of concentration and other aspects of English learning in VR between environments with and without built-in feedback functions and to determine which environment is better able to promote a flow state, thus informing the future design of environments that may enhance learning opportunities.

Participants

This experiment was conducted at a science and engineering university in Saitama, Japan. 11 pairs of participants were created (thus, 22 participants). The participants ranged in age from 20 to 27, with an average age of 21.6. All participants were Japanese.

Environment overview

Game-like environments with and without built-in feedback functions were constructed to measure flow. The equipment and software used to develop these environments and the details of the construction of each are described below.

Equipment and software used

The following equipment and software were used in this study.

1) Head-mounted display (HMD) A head-mounted display (HMD) is a display device worn on the head and capable of projecting three-dimensional images by utilizing the parallax between the left and right eyes. In this study, we used the Vive (Figure 1). This is an HMD for VR jointly developed by HTC and Valve. This HMD is equipped with two motion-tracking sensors that determine the user's exact position in space, allowing the user to move around and not just sit down. In this study, we controlled the movement of the avatar in the VR space by using the positional information of the HMD and controllers.

Figure 1

The HTC Vive head-mounted display, controllers, and sensors.



2) Game production engine Unity, developed by Unity Technologies, was used as the game production engine for this study. Unity is a game engine that can be used on a variety of platforms, including mobile, desktop, game consoles, and the web. Since it was originally developed for 3D games, it was easy to create VR environments for the study.

3) Voice call software In this study, Discord developed by Discord, Inc. was used as the voice call software.

4) 3D models Liam, created by AKISHAQS, was used as the 3D avatar. It is a model of an adult male without facial expressions and can be used for animation, games, and VR/AR projects. Additional assets were downloaded from the Unity Asset Store.

Environment contents

Two VR environments were created for this study; in both environments, pairs of participants performed a spot-the-difference task. In each environment, participants were presented with a dollhouse in which a variety of objects were placed (Figure 2). Some of the objects in each room were placed in the same location, while others were placed in different locations. The environment was designed so that through communicating with each other participants find some objects in different locations and, work together to place them in the same location within their respective dollhouses. Five levels of difficulty were created. Level difficulty was operationalized via the number of objects in different locations (in terms of Robinson's framework: +/- number of elements to be manipulated).

Figure 2

Dollhouse presented to each participant in a dyad.



Participants can see their own room, but cannot see the contents of their interlocutor's room. This was achieved by placing the rooms separately and of opposite orientations to each other (see Figure 3).

Figure 3

Subject's point of view showing that their interlocutor's dollhouse content is obscured.



Wearing an HMD enables an immersive VR experience. participants can move their heads to understand what is going on in the room and gesture to others by moving the controller (Figure 4).

Figure 4

Experiencing the VR environment



Of the three elements considered essential for promoting flow, this experiment is

concerned with manipulating the second: feedback. Thus, the experiment utilized two versions of the same VR environment: one with feedback and one without. The specific differences between the two environments are described below.

The environment with feedback

Both visual and audio feedback was added to this environment, making it more game-like than the no-feedback version. Visual feedback was provided via a score (10 points per object placed in a correct position) and elapsed time for each level displayed in the subject's field of vision. In addition, audio feedback was utilized. First, a sound effect plays when participants place objects in the same location, and secondly, a different sound effect plays when participants complete a level. These effects indicate whether participants have made a correct or incorrect decision and whether the task is completed (Figure. 5).

Figure 5

Environment with feedback



The environment without feedback

In the environment without feedback, participants do not receive a score or see their elapsed time, and instructions to proceed to the next level are displayed only when the level is completed. In addition, no sound is played when an object is placed in the same location, meaning that participants must confirm the location of objects orally (Figure 6).

Figure 6

Environment without feedback

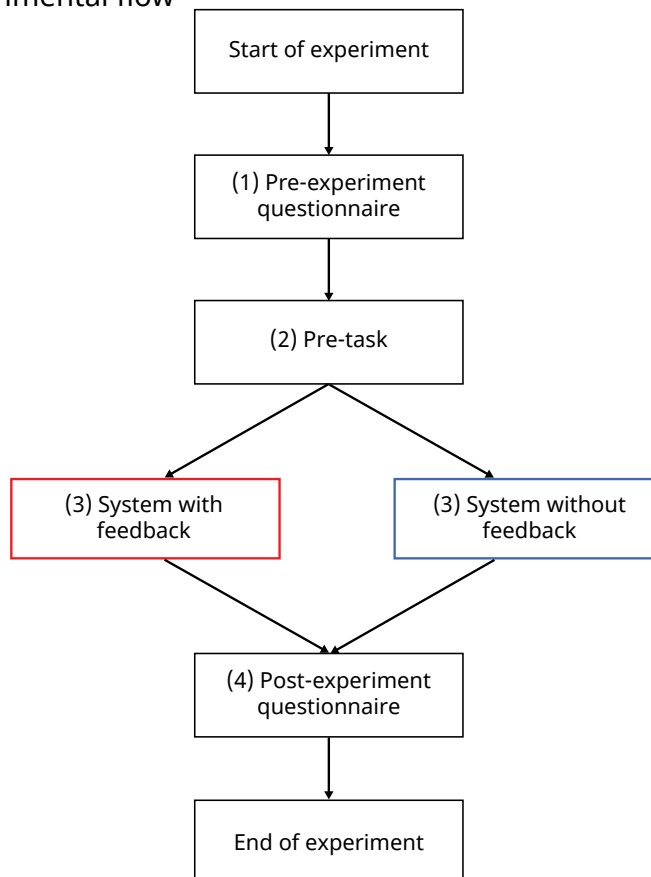


Experimental procedure

The experiment followed the flow shown in Figure 7 below.

Figure 7

Overall experimental flow



(1) Preliminary Questionnaire

Before performing the task, a questionnaire was administered to determine participant demographics such as age and gender.

(2) Pretask

A pre-task worksheet was employed to prime the participants to use vocabulary and prepositions which appear within the environment. Participants completed cloze questions which used the same graphics as those within the environment as an additional way to acclimatize participants to the upcoming communication task. The pre-task was employed to minimize the effect of English proficiency on oral performance.

(3) Task cycle for both environments

Each task cycle was performed as follows.

- 1 Operation Confirmation:** Participants were given time to acclimatize to the environment and test the HMD, microphones, and earphones.
- 2 Spot-the-difference Task:** Participants completed five levels within one of the two environments.

(4) Post-assessment questionnaire

A questionnaire was used to measure flow. Details are described below. In order to examine the sense of time distortion as found with a flow state, participants were prompted to indicate the amount of time they experienced from the start to the end of the activity.

Questionnaire content

Participants were asked to complete a questionnaire to investigate the degree of their flow state after experiencing the spot-the-difference activity (Table 1). The questionnaire was based on Cho (2018) and is intended to uncover which of three constructs (interest, attention, and control) were positively affected during an activity. Ratings on the items were measured on a 6-point scale to avoid central effects.

Results and discussion

RQ1: Flow state data

The questionnaire was categorized into three items, interest, attention, and control, which are considered the fundamental elements to promote a flow state. To these three items, we added one additional item, the flow score, which is the average score for all three elements. This gave a total of four items. We analyzed the questionnaire data using an unpaired t-test. Results are shown in Table 2.

Table 1

Statements in the post-experiment questionnaire (based on Cho, 2018)

N°	Flow measurement	Statement
1	Interest	I would do this task even if it were not required.
2		This task was interesting in itself.
3		I found the experience very rewarding and felt good after completing it.
4		This task aroused my imagination.
5	Attention	It took no effort to keep my mind on the task.
6#		When doing this task, I was aware of distractions.
7		When doing this task, I was totally absorbed in what I was doing.
8	Control	When doing this task, I knew clearly what I wanted to do.
9		When doing this task, I had a feeling of control over what and how to write or speak.
10		When doing this task, I had a feeling of total control.

*Item 6# was reversed coded.

Table 2

T-test results for the four items

	With feedback		Without feedback		Mean difference	p-value	sig.
	Mean score	SD	Mean score	SD			
Interest	5.4	0.57	4.9	0.62	0.5	0.092	*
Attention	5.5	0.59	5.4	0.64	0.1	0.79	
Control	4.7	0.78	4.6	1.04	0.1	0.93	
Flow	5.2	0.54	5.0	0.60	0.2	0.46	

* 0.05 < p < 0.1

Mean scores for the variable “flow” were higher than the mean of 3 for both environments (with feedback = 5.2, without feedback 5.0). However, there was no statistically significant difference between mean scores for the two environments for this measure ($p = 0.46$). This indicates that both environments were able to promote a flow state in participants, but that feedback did not play a significant role in promoting flow. Of the three individual elements (interest, attention, and control), there was a statistically significant difference between the mean scores for the Interest component ($p < 0.01$). This indicates that the provision of feedback

increased participants' level of interest in completing the language task. No statistically significant differences were found for the remaining two components.

Subsequently, individual items on the questionnaire were examined for potentially significant differences in mean scores between the two environments. Analysis was conducted using unpaired t-tests for each item. Results are provided in Table 3.

Table 3

Statistical results for each item on the flow questionnaire.

Item	With feedback		Without feedback		Mean diff.	p	Sig.
	Mean	SD	Mean	SD			
1. I would do this task even if it were not required.	5.08	1.00	4.60	0.97	0.48	0.26	
2. This task was interesting	5.92	0.29	5.50	0.53	0.42	0.03	**
3. I found the experience very rewarding and felt good after completing it.	5.92	0.29	5.10	0.88	0.82	0.01	**
4. This task aroused my imagination.	4.58	1.08	4.50	0.85	0.08	0.85	
5. It took no effort to keep my mind on the task	5.42	0.67	5.50	0.71	-0.08	0.78	
6#. When doing this task, I was aware of distractions.	5.42	0.79	5.10	1.45	0.32	0.52	
7. When doing this task, I was totally absorbed in what I was doing	5.58	0.51	5.60	0.70	-0.02	0.95	
8. When doing this task, I knew clearly what I wanted to do	5.67	0.49	5.00	1.49	0.67	0.16	
9. When doing this task, I had a feeling of control of what and how to write or speak	4.50	1.31	4.50	1.65	0.00	1.00	
10. When doing this task, I had a feeling of total control	3.83	1.11	4.40	1.17	-0.57	0.26	

Results for two items in the Interest category (“This task was interesting” and “This task was challenging and gave me a sense of accomplishment”) were significantly higher when completing the task with feedback.

RQ2: Comparison of elapsed time and perceived time of completion

Participants were asked to respond to an item regarding their perception of elapsed time. These perceptions were then compared with the actual elapsed time in order to measure one of flow's main characteristics: a distortion regarding the sense of elapsed time during an activity. The environment with feedback was completed in a shorter time than the environment without feedback, and perceptions for this domain were also shorter. However, there was no significant difference found between mean scores (Table 4). Of note is that the elapsed time and perceived time for the task with feedback were shorter than the task without feedback. This is explored later in the discussion section.

Table 4

Mean (and standard deviation) of perceived time and elapsed time for both VR environments.

Environment	Perceived time	Elapsed time	Mean diff.	p-value	sig.
With feedback	7.0 (3.2)	7.1 (0.9)	0.1	0.92	
Without feedback	11.2 (4.3)	10.3 (1.8)	-0.9	0.57	

RQ3: Participant perceptions of the environments

Participants were also asked to respond to an open-ended question regarding their opinions of the VR environments. Comments are presented in Table 5 (translated from Japanese). Positive comments suggest that participants enjoyed learning with the VR environment regardless of the presence or absence of feedback. Enjoyment was promoted due to the immersive experience provided by the VR, and the sense of accomplishment in completing the tasks through physical movement. Negatively weighted comments mentioned that participants were confused about whether they were in charge of the questioning or answering and that they had trouble when they could not remember task-critical vocabulary.

Table 5
Comments from participants

Participant	Comment
With_01	It was my first time using VR and I think it was very enjoyable and educational.
With_02	It was fun.
With_03	I might have enjoyed studying English more if I had used this environment when I began to learn it.
Without_01	I was a little worried that I would bump into the desks and walls around me, but I was quite immersed in the work and enjoyed the activity without worrying about my surroundings.
Without_02	It was interesting. I thought that unifying the categories of nouns (animals, tools, etc.) would make it easier to understand and learn in a unified manner.
Without_03	Sometimes I'm not sure if I should be the one asking the questions or the one answering them.
Without_04	The task was a little easy and it seemed very doable with more variation. I also felt that a word cheat system would be useful for when I got stuck on a word.
Without_05	I was quite excited by my first VR game experience. I felt a sense of accomplishment in moving my body to complete tasks.
Without_06	It was easy to speak English because the group was small.

Discussion and conclusion

Data obtained from the post-experiment questionnaires revealed no significant difference between the two environments in terms of their ability to generate a flow state in participants. One reason for this may be that unlike in studies like Cho (2018) and York et al. (2021) where participants experienced multiple different environments, in this study each participant experienced only one environment, meaning that they were unable to compare their experiences to another mode. Therefore, having participants perform tasks in both environments may produce results of improved accuracy.

The flow questionnaire revealed statistically significantly higher results for two items on the interest measure when feedback was provided: that the task was more interesting and that it was more rewarding and provided a sense of accomplishment. Informal observations of participants completing the spot-the-difference tasks revealed that when feedback was not provided, they struggled to understand whether an object was correctly placed which created a state of

confusion regarding how they should proceed. This was alleviated with the environment which provided feedback as an audio cue sounded to indicate that an object was in the “correct” location prompting participants to move on to the next object. Based on these observations, an improved sense of control may be developed when feedback is provided, however, results did not show a significant difference in mean scores for control. Thus, further exploration is required.

Regarding the comparison of perceived and actual elapsed time, the results showed that both elapsed time and perceived time were shorter with feedback than without. The reason for this result may be that the task was easier to understand with feedback than without since the results of the activity were immediately apparent and there was less unnecessary confusion. Additionally, as part of the environment with feedback, a timer was provided on the screen to the participant so they could see the elapsed time of each level they completed. Thus, participants could have calculated the elapsed time by paying attention to the time they spent on each level. However, this was not explored in any significant way in this study.

Finally, participants’ comments regarding the two environments revealed that it was generally enjoyable, but this could be attributed to a novelty effect. Indeed, one comment (With_01) mentioned explicitly that this was their first time experiencing VR. Of similar note is a negative comment from a participant that experienced the environment without feedback (Without_03). They mentioned that they were unsure whether they should be asking or answering questions. This comment connects to the feeling of confusion that was promoted by the environment without feedback.

Future research

The results of the present study showed significant differences in two items of Interest. However, due to conducting this experiment during the COVID pandemic, it was not possible to gather a sufficient number of participants to explore the generalizability of results further. In addition to the items for which significant differences were found, there were other items for which significant trends were found, and we believe that future experiments with a larger number of participants may yield more precise results.

Feedback in this study was operationalized in terms of audio feedback and the addition of a timer for each level. Exploration of other forms of feedback such as visual and tactile/haptic may help to further immerse learners in the VR domain and thus an improved state of flow. Additionally, as mentioned by participant *Without_04*, the difficulty-skill balance of the tasks may not have been calibrated well enough for the participants meaning that tasks were too easy. Task difficulty could be operationalized via the environment in a future study to ensure that the challenging nature of the tasks was optimized for the participants level of skill. This could be achieved by increasing or decreasing the number of elements that participants had to manipulate based on how quickly they completed a task.

References

- Cho, M. (2018). Task Complexity and Modality: Exploring Learners' Experience From the Perspective of Flow. *The Modern Language Journal*, 102(1), 162–180. <https://doi.org/10.1111/modl.12460>
- Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement*. Springer
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper and Row.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Collins.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed). Longman/Pearson.
- Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, 32(3), 158–172. <https://doi.org/10.1007/s11031-008-9102-4>
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179. <https://doi.org/10.1016/j.chb.2015.07.045>
- Kim, D., & Ko, Y.J. (2019). The impact of virtual reality (VR) technology on sport spectators' flow experience and satisfaction. *Computers in Human Behavior*, 93, 346–356.
- Kormos, J. (2014). Differences across modalities of performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes & R.M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 193–216). John Benjamins
- Ministry of Education, Culture, Sports, Science, and Technology (MEXT). (2014). *Report on measures to improve and enhance English language education in the future - Five recommendations for reforming English language education in response to globalisation*. Available at: https://www.mext.go.jp/b_menu/shingi/chousa/shotou/102/houkoku/attach/1352464.htm
- Nakamura, J., & Csikszentmihalyi, M. (2014). The Concept of Flow. In M. Csikszentmihalyi (Ed.), *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (pp. 239–263). Springer. https://doi.org/10.1007/978-94-017-9088-8_16.
- Perttula, A., Kiili, K., Lindstedt, A., & Tuomi, P. (2017). Flow experience in game based learning – a systematic literature review. *International Journal of Serious Games*, 4(1), Article 1. <https://doi.org/10.17083/ijsg.v4i1.151>
- Robinson, P. (2007). Criteria for grading and sequencing pedagogic tasks. In M. del Pilar Garcia Mayo (Ed.). *Investigating Tasks in Formal Language Learning* (pp. 7–27). Multilingual Matters.

- Rutrecht, H., Wittmann, M., Khoshnoud, S., & Igarzábal, F. A. (2021). Time speeds up during flow states: A study in virtual reality with the video game thumper. *Timing & Time Perception*, 9(4), 353–376.
- Shibata, K., & York, J. (2021). A comparison of the affective affordances of a static and interactive VR system on learner FLA and motivation. In E. Forsythe (Ed.), *Teaching with Tech* (pp.108-127). JALT CALL SIG, Japan. <https://doi.org/10.37546/jaltsig.call2020.8>.
- Tokutake, H., York, J., Shibata, K. & Nakayama, H. (2021). Evaluation of a VR Language Learning System: Effect of Interactivity on Learner Oral Task Performance. *Journal of Educational Multimedia and Hypermedia*, 30(2), 183–203. Waynesville, NC USA: Association for the Advancement of Computing in Education (AACE). Retrieved October 26, 2022 from <https://www.learntechlib.org/primary/p/219046/>.
- York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2021). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL*, 33(1), 49–70. <https://www.doi.org/10.1017/S0958344020000154>.

Authors' bios

James York is a senior assistant professor at Meiji University where he conducts research on ludic language pedagogy.

Hirotaka Takuma is a masters student at Tokyo Denki University. He is exploring research on the use of virtual reality in language learning contexts.

Takuya Yoshino is a masters student at Tokyo Denki University. He is exploring research on the use of virtual reality in language learning contexts.

Hiroshi Nakayama is a professor of Information Systems Design at Tokyo Denki University, Faculty of Science and Technology. He conducts research on instructional design and educational technology.

ISBN 978-4-901352-70-3

Playful CALL
Selected papers from the JALTCALL2022 Conference



JALTCALL SIG <https://jaltcall.org>