Productive Failure in Virtual Language Learning for English

By

Puji Rahayu

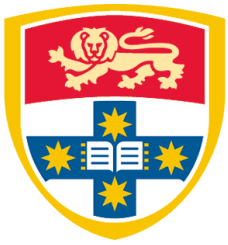
A thesis submitted in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Sydney School of Education and Social Work

The University of Sydney

August 2021



THE UNIVERSITY OF SYDNEY

DECLARATION

This is to certify that:

(I) this thesis comprises only my original work towards the degree of Doctor of Philosophy

(II) due acknowledgement has been made in the text to other materials used

(III) this thesis does not exceed the word limit for this degree

(IV) no part of this work has been used for the award of another degree

(V) this thesis meets the University of Sydney's Human Research Ethics Committee (HREC) requirements for the conduct of this research

Statement of use of a third party for editorial help

I hereby confirm that this thesis was proofread and edited for the language, spelling, and grammar by Abhirama S. D.

Puji Rahayu

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Vocabulary and syntax are a challenge for English as Foreign Language (EFL) learners when they want to communicate in English. Task-based Language Teaching is commonly used in EFL teaching of vocabulary and syntax, which is a type of Direct Instruction (DI) that involves the initial use of explicit language instruction followed by a language learning activity. This study compared the efficacy for language learning of a different type of pedagogical approach, Productive Failure (PF), which delays instruction until after a language learning activity, to Direct Instruction (DI). There were three main language learning assessment areas: (a) students' declarative and procedural knowledge in the written production of the target language, (b) students' declarative and procedural knowledge in the spoken production of the target language, and (c) students' cognitive and metacognitive strategies in learning. English language education department freshmen in an Indonesian university (N=112) participated in the study by performing language learning activities in Second Life (*SL*), which is a 3-D virtual learning environment. They were randomly assigned to two language learning treatment groups. The PF group finished a communicative task on describing places prior to receiving explicit instruction. In contrast, the DI group watched an instructional video before completing a communicative task on describing places. This was followed by students in both groups finishing a similar communicative task in *SL*. Data from pre-and post-tests were analysed quantitatively, and video captures were transcribed and analysed qualitatively. The quantitative results found that PF group students performed significantly higher on the English syntax written assessment and both groups performed equally on the written vocabulary assessment. However, both groups performed equally on the spoken assessments of syntax and vocabulary. In the qualitative analysis, the PF students were found to use more self-regulated learning strategies and study tactics than DI students. The pattern of these findings is discussed in terms of previous research and theory. Overall, these findings suggest further research is warranted to investigate the use of PF language learning activities that involve the use of a virtual learning environment.

Key words:

Productive failure, MUVE-based EFL learning, Self-regulated learning, procedural knowledge, task sequencing

ACKNOWLEDGEMENT

Alhamdulillah, all praises to Allah, the Most Gracious, most knowledgeable (sustaining)

This PhD journey has been *kawah candra dimuka* – a training place – for me as an academics, a woman, and a mother of two. I owed countless gratitude to many parties that have supported the long-winding journey to finish this research. That gratitude goes to:

1. my supervisor, Prof. Michael Jacobson, my sincerest thanks for the support during this journey, which was sometimes beyond my capability. Your motivation through your encouraging comments and trust boosted my energy to finish this PhD. More importantly, your patience in teaching me to always learn the research skills. Also, I would like to thank Asc. Prof. David Hirsh for his helpful comments and books on language teaching & learning. You have all contributed to my development as a researcher. Sincere appreciation goes to both of you, who have always motivated and encouraged me over my self-doubt by comforting me and assuring that my research does contribute to both theory and practice.
2. the Indonesian government, via the Endowment Fund for Education, *Lembaga Pengelola Dana Pendidikan* (LPDP) Indonesia for their generous sponsorship of PhD. I am also thankful to Universitas Islam Indonesia, Yogyakarta, for the permission and financial support during my additional time beyond the standard study period. Without the supports from both institutions, I would never imagine having such experience of pursuing my PhD at the University of Sydney.
3. my colleagues at Centre for Research on Learning and Innovations (CRLI), Sydney School of Education and Social Works, The University of Sydney, for the generous academic and life sharing. I always enjoyed the discussions, evening talks, and the strolls we had. I am also thankful to Indonesian friends who have been my family when my immediate family was miles away. Special thanks go to the PhD moms—Arum, Ayu, Feti, Ijan—who always offered their hands when I needed helps. Thanks to family in the "LPDP Sydney Uni" and "RW Campsie" for the experience-sharing during the hard times in this PhD life. I also thank my spiritual teacher—Uya—for always reminding me to be strong in this challenging condition. May Allah give you *barakah* in your life.
4. my friends, Ditya, Rosa, and Anita for their valuable feedbacks. They have invested their time reading the final draft of this dissertation despite their limited time.
5. my kids Ihsan, Iqbal, and my husband Kuncoro for their supports, both direct and indirect. Ihsan and Iqbal have been significant parts of this dissertation. I could not imagine finishing this long and winding PhD journey without them beside me. My husband, Kuncoro, has given me more opportunities to be a stronger PhD mom in his unique way.
6. my mom, Mujiyem, and my late father, Surasa, who have brought me up to become a strong woman. My late brother, Suwarta, who passed away when I was miles away.
7. my brother, Wuryanta, and my sister, Sri Asih, and their family who have supported me sincerely.

I thank you all for all your sincere supports.

ABBREVIATIONS

CAF	complexity, Accuracy, and Fluency
CALL	computer-assisted Language Learning
CBLE	Computer-based learning environment
CCS	Complex-complex-simple
COPES	Conditions, Operations, Products, Evaluations, and Standards
CSC	Complex-simple-complex
DGBL	Digital game-based learning
DI	Direct instruction
DIY	Daerah Istimewa Yogyakarta
EFL	English as a Foreign Language
ELE	English Language Education
FLT	Foreign language teaching
FoK	Feeling of Knowing
HL	High-Low
JoL	Judgement of Learning
LH	Low-High
LPDP	*Lembaga Pengelola Dana Pendidikan,* the main scholarship provider
LRE	language-related episodes
MUVE	Multi-user Virtual Environment
PBI	Production-based Instruction
PF	Productive Failure
PPP	Presentation, Practice, Production
RSMs	representation and solution methods
SCS	simple-complex-simple
SL	Second Life
SRL	Self-regulated Learning
SSARC	Stabilise, Simplify, Automatise, Reconstruct, and Complexify
SSC	Simple-simple-complex
STEM	science, technology, engineering, and math
TBLT	Task-based Language Teaching

Chapter 1: Research Overview

1.1. Introduction

Effective guidance has been the topic of interest among researchers on pedagogical strategy. The discussion covers the degree of guidance—high or low—and sequence of guidance—before or after problem-solving (Lai, Portolese, & Jacobson, 2017). For the degree of guidance, some experts, on one hand, posit that high guidance will be useful for learners as it helps learners receive the critical concepts and procedures to learn and is cognisant to human cognitive architecture (Kirschner, Sweller, & Clark, 2006). Examples of high guidance include lectures, demonstration, presentation, and guided practice. On the other hand, low guidance enthusiasts claim that learning is effective when guidance is limited. Students need to discover or construct critical concepts by themselves for learning to be effective and meaningful (Kirschner, Sweller, & Clark, 2006). Examples of low guided learning can be seen in discovery learning (Mayer, 2004), inquiry learning (Hmelo-Silver, Duncan, & Chinn, 2007), problem-based learning (Hmelo-silver, 2004), or experiential learning (Kolb, 1984, 2014). Alternatives to merge high and low guidance are also presented.

A combination of low and high guidance comprises high-low guidance (e.g., DI or Apprenticeship) and low-high guidance of learning (e.g., Productive Failure), which is related to the sequence of guidance. In high-low guided learning, students receive instruction or demonstration (high guidance) at the beginning of the lesson for the acquisition of key concepts and procedures (Kirschner et al., 2006; Kirschner & Clark, 2004), then they do guided practices (medium/low guidance). The guidance fades away gradually until students are ready for a problem-solving activity (no guidance). In contrast, low-high guided learning lets students solve a complex problem without helps at the beginning followed by an instruction or demonstration to acquire key concepts and procedures (Jacobson, Kim, Pathak, & Zhang, 2015; Kapur & Bielaczyc, 2012). The purpose of the problem-solving activity is to activate prior knowledge and acknowledge gaps, which will then be filled in the subsequent instruction (Loibl & Rummel, 2014a; Newman & DeCaro, 2019). Which combination is used in English language teaching?

In English as a Foreign Language (EFL) learning, the tasks are commonly sequenced in the simple-complex task sequence (high-low guidance), which is similar to the Presentation, Practice, and Production (PPP). In PPP (Cook, 2008), an example of the weak-form application of Task-based Language Teaching (TBLT), materials are presented at the beginning of the lesson (simple task, high guidance), followed by guided practices (medium

task, medium guidance), and more authentic use of the target language in the production stage (complex task, low guidance). This sequence is similar to the Stabilise, Simplify, Automatise, Reconstruct, and Complexify (SSARC) model of task sequencing, which suggests simple-complex task sequencing (Robinson, 2010, 2015) as the most effective sequence. Simple tasks help stabilise the newly acquired knowledge (lexical and grammatical forms); medium tasks give more opportunities for similar idea-expression; and complex tasks pull students' potential for form-meaning connection in their spontaneous speech act (Allaw & McDonough, 2019; Lambert & Robinson, 2014; Robinson, 2006, 2015). Both simple-complex task sequence and PPP prescribe an explicit instruction prior to the simple task to acquire new knowledge. This new knowledge is practised in the simple and medium tasks until it can be used entirely without guidance in the complex task.

While studies reported the success of TBLT (Kim, 2020; Wang, Petrina, & Feng, 2017), it is also evidenced that students are unwilling to speak the target language in the classroom (Carless, 2002; Criado, 2013). To the author's observation, students are reluctant to speak in the target language because they can in fact communicate in the mother tongue. They do not have authentic reasons to use the target language in the class, which is one of the reasons for successful spoken-based instruction (Ozverir, Herrington, & Osam, 2016). In addition, a study carried out in Indonesia found that one of the reasons for students' unwillingness to communicate in the target language is inadequate vocabulary and syntax knowledge to speak (Cahyono, 2008). As a result, solutions for the inauthenticity and inadequate vocabulary need consideration.

Immersing students in the target language countries is one of the potential solutions. Studies on language immersion program revealed better achievement in language proficiency and skills in immersion students than non-immersion classes (Cheng, Li, Kirby, Qiang, & Wade-Woolley, 2010; Knell et al., 2007; Rugasken & Harris, 2009; Savage & Hughes, 2014). For example, in a mixed-method approach, 140 US cadets participated a Chinese summer language program in Nanjing University, China. They attended a 20-hour classroom-based course each week focusing on listening, speaking, reading, and writing. They also joined in cultural excursions every weekend. A pre-and post-tests were assigned before and after the program. The results showed that improvements were recorded in the listening and reading. Participants claimed it was due to their authentic communication with the first speakers of the target language (Rugasken & Harris, 2009). Further investigation suggested that the authenticity of the interaction had a significant influence on an immersion program (e.g. Borreguero Zuloaga & De Marco, 2020; Savage & Hughes, 2014).

While immersion in the target country can help students learn a foreign language, it may not be practical in terms of budget and logistical constraints. Immersing students in an English-speaking country can be very expensive for low-income students. For example, one student staying in Melbourne for two weeks for an immersion program can spend more than 30 million Rupiah, equal to 10-month local Indonesian income. The budget constraints will likely entail in logistical problems. Therefore, alternatives to the physical immersion are needed for the non-immersion program to succeed.

As an alternative to *the physical immersion* of students in an English-speaking country, there have been many studies exploring how Multi-user Virtual Environments (MUVEs) might provide *virtual immersion* for students to learn English and other foreign languages (Hong, Jeong, Kalay, Jung, & Lee, 2016; Nelson, Ketelhut, Clarke, Bowman, & Dede, 2005; Peterson, 2011; Rahayu & Jacobson, 2012; Wang, 2015). The studies suggest that unlike classroom-based language learning, MUVE-based language learning encourages students to be more willing to speak in the target language, use more hands-on spoken discourse, and apply longer utterances in speaking. For example, one of the participants helped a friend change appearance in Second Life (*SL*). He described how to buy, take off, and alter the clothes in both voice and text messages (Rahayu & Jacobson, 2012). In the interview, he commented he felt confident that his spoken English was understood, which suggested he had speaking self-efficacy (Bandura, 1997). He also felt the genuine need to help his friend to use any expression he needed to convey his intended meaning (for detail, see Rahayu & Jacobson, 2012). Other studies show that learning languages in MUVEs increases self-efficacy (Huang, Grant, & Henderson, 2012; Rahayu & Jacobson, 2012; Zheng, Young, Brewer, & Wagner, 2009), social interaction (Marklund, Backlund, Dahlin, Engström, & Wilhelmsson, 2014), and supports authentic communication experience (Lan, Kan, Hsiao, Yang, & Chang, 2013; Rahayu & Jacobson, 2012).

Despite the reported benefits of using MUVEs in EFL learning, there is a lack of research on effective pedagogical approaches in MUVE-based EFL learning. There are two main options for research that would explore this issue of identifying effective pedagogies for EFL involving MUVE systems: (a) employ existing EFL pedagogical approaches, and (b) employ innovative pedagogical approaches used in other subject areas. An example of the first approach is the PPP technique (Cook, 2008) discussed above, which could be implemented in an EFL-MUVE system. In contrast, an example of the second approach is a recent learning design called PF (Kapur & Bielaczyc, 2012), which has received considerable research attention for learning in scientific and mathematical areas (Kapur, 2008, 2014b,

2016). This approach has also been successfully implemented in a virtual learning environment (similar to a MUVE) to help students learn about scientific inquiry (Jacobson, Taylor, & Richards, 2016).

This research investigated these two above mentioned pedagogical approaches (PF and DI) for learning EFL in a MUVE. The rationale for this selection is the apparent cognitive theoretical differences between the two different approaches. DI is effective due to the *reduction of cognitive load* that would help students learn correct knowledge and procedures better and reduces misconceptions and misunderstanding (Chandler & Sweller, 1991). In PPP techniques, presentation (P1) provides explicit instruction about linguistic items and structures. Practice (P2) ensures the mastery of linguistic items and structures in guided activities like gap filling. Once the linguistics items and structures are mastered, production (P3) activities are assigned for the students, in which no guidance is provided (Criado, 2013). Kirschner and colleagues (2006) maintain that DI approaches lead to more significant learning gains at introductory and intermediate levels, in comparison to other pedagogical approaches.

In contrast to DI, PF learning designs have a very different theoretical grounding, such as the *activation of prior knowledg*e, *knowledge gap recognition,* and *schema abstraction* (Jacobson et al., 2020; Kapur, 2008). Students' prior knowledge is activated and differentiated during the initial problem-solving activities as they can only rely on their prior knowledge. When prior knowledge is activated, the general knowledge gap is potentially acknowledged (Loibl & Rummel, 2014a). Additionally, the specific knowledge gap is admitted by the activity of *comparing and contrasting* students' solutions to the canonical or "expert-like" solutions (Kapur, 2008; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a). Recent PF-related studies have found that PF leads to significantly higher learning outcomes than DI approaches in mathematics, statistics, and science topics, particularly concerning the application or use of knowledge to solve novel problems, i.e., knowledge transfer (Jacobson et al., 2020, 2017; Kapur, 2014b; Lai et al., 2017; Newman & DeCaro, 2019)

The above mentioned research findings that PF can better lead to enhanced learning than DI pedagogical approaches have potential implications for EFL teaching. Therefore, this research's primary purpose was to compare PF to DI – as the commonly applied pedagogical strategies in EFL teaching – in MUVE-based English learning activities. The research focused on the effect of both strategies on students' syntactic and lexical accuracy in written and spoken description of places. Besides, this study aimed to compare students' self-regulated learning (SRL) strategies between PF and DI provided that studies on SRL revealed

that prior knowledge – one of the PF mechanisms – was one of SRL moderating factors that affected an accurate goal-setting (Moos & Azevedo, 2008a; Taub & Azevedo, 2017; Taub, Azevedo, Bouchet, & Khosravifar, 2014; Trevors, Duffy, & Azevedo, 2014).

1.2. Aims and Research Questions

The study compared the efficacy of two different pedagogical strategies – PF and DI – in learning vocabulary and syntax of written and spoken performance, especially in describing places. This study utilised three indicators to measure the impact of the two pedagogical strategies: 1) students' declarative and procedural knowledge in the written production of the target language, 2) students' declarative and procedural knowledge in the spoken production of the target language, and 3) students' cognitive and metacognitive strategies in learning, especially during the explicit instruction in the consolidation phase. The research questions of the study are:

1. Is there a different achievement in the written description of places between the experimental (PF) and control (DI) groups?
    a. Overall written description
    b. Procedural knowledge
    c. Declarative knowledge
2. Is there a different achievement in the oral description of places between the experimental (PF) and control (DI) groups?
    a. Overall written description
    b. Procedural knowledge
    c. Declarative knowledge
3. Are there different behaviours during the instructional video between the experimental (PF) and control (DI) groups from self-regulated learning theorising?

In line with the results of the previous studies of PF (Jacobson et al., 2017; Kapur, 2012; Kapur & Bielaczyc, 2012; Loehr, Fyfe, & Rittle-Johnson, 2014), the hypothesis of the first question is that there is equal learning of vocabulary (declarative knowledge) between experimental and control groups, but there is a difference in the learning of syntax (procedural knowledge) between the experimental and control groups. This hypothesis was generated based on vocabulary being declarative knowledge, while syntax is a procedural one (Paradis, 2009).

In addressing the second research question, the hypothesis put forward is that both PF and DI groups use an equal number of vocabulary (declarative knowledge) in the spoken description of places. Yet, DI students use more syntax (procedural knowledge) in the

description of places compared to the PF students. Previous studies on language learning have found that post-instruction practices play an essential role in using newly-learned expressions in an oral production of the target language (e.g., Finardi, 2008; Fukuta, 2016; Ho, 2017; Jacoby, 1978). In the PF-related studies, it was found that the number of practices supported procedural fluency (used interchangeably with procedural knowledge) after instruction provision on the canonical problem solving (Rittle-Johnson, Siegler, & Alibali, 2001; Sleeman, Kelly, Martinak, Ward, & Moore, 1989). In the context of this study, considering DI students had more practices after the explicit instruction, they should use the newly learned vocabulary and syntax in the oral production of describing places better.

For the third research question, the hypothesis is that PF students generate more meta-cognitive monitoring and cognitive learning strategies. As PF students activated more prior knowledge during the first problem solving of the PF pedagogical strategy (Kapur, 2008; Kapur & Bielaczyc, 2012), they should set their sub-goals more accurately. In addition, comparing students' solution and the canonical solution in the consolidation phase facilitates knowledge gap recognition (Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a). When sub-goals are accurate, and knowledge gaps are recognised, it is easier to choose appropriate study tactics during the instruction. In addition, the canonical solution presented in the consolidation functions as the standard to monitor the learning strategies used. Therefore, it is reasonable to predict that PF students will enact more self-regulatory strategies than DI students.

1.3. Significance of the Study

This study is significant in four main ways. First, the study is one of the first to explore the use of PF learning design in language learning, especially around its effect on the vocabulary and syntax of a spoken and written description of places. The learning mechanisms in PF pedagogical strategies have proven its superiority over DI on the procedural fluency, conceptual knowledge, and transfer have been carried out in STEM-related areas (e.g., Ashman, Kalyuga, & Sweller, 2020; Kapur, 2014b, 2015; Lai et al., 2017; Loibl & Rummel, 2014b). These benefits could be tried out in foreign language teaching as a similar strategy has been applied in the EFL context. While a complex-simple task sequence used in EFL is similar to the PF design, the position and the design of the explicit instruction is different (Allaw & McDonough, 2019; Robinson, 2010, 2015). While the explicit instruction in the simple-complex task sequence in foreign language teaching (FLT) is positioned before the first simple task (Robinson, 2015), it is prescribed to follow a complex task – a novel-problem solving – in PF (Kapur & Bielaczyc, 2011). Besides, there is no

specific method of the instruction in the simple-complex task sequence in FLT, yet a comparison and contrast between students' solution – as the product of the initial problem solving – and the canonical solution is prescribed in the PF design (Kapur & Bielaczyc, 2011). Therefore, applying PF in EFL teaching will contribute to the alternative design on the task sequencing in the application of TBLT.

Second, this study used communicative tasks as both a tool to activate prior knowledge for future learning preparation (Schwartz & Bransford, 1998) and practice the newly-learned knowledge from the instruction and schema abstraction (Jacobson et al., 2020; Kapur & Bielaczyc, 2012). In a foreign language teaching context, in contrast, "authentic"[1] communicative tasks function as the final product for students to master, either in the simple-complex (Allaw & McDonough, 2019) or complex-simple task sequences (Baralt, 2014) in the weak-form application of TBLT. PF-related studies have proven the benefits of prior knowledge activation and knowledge gap recognition as a preparation for the effective future learning in the subsequent instruction (Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), yet no studies have been carried out in the EFL context. Applying the PF design is expected to add the explicit instruction's effectiveness in the weak-form application of TBLT in terms of understanding the cognitive mechanism of learning.

Third, understanding the comparison between PF and DI in terms of cognitive learning and metacognitive monitoring strategies used in the explicit instruction contribute to the effect of both strategies on self-regulated learning. Previous PF-related studies have found the benefits of prior knowledge activation and differentiation and global and individual knowledge gap recognition as one of the moderating factors of successful learning from a subsequent instruction (e.g., Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014b, 2015). Provided that these PF-related learning mechanisms – such as prior knowledge – are related to self-regulated learning (e.g., Moos & Azevedo, 2008a; Taub & Azevedo, 2017), this study would give empirical evidence on the differential effects of PF and DI on self-regulated learning strategies during the explicit instruction.

Fourth, this study brought about significant practical implications for MUVE-based learning programs. As this study employed an innovative virtual environment called Second Life with an empirically tested pedagogical strategy – PF – to affect student's vocabulary and syntax in describing places, it could be an alternative pedagogy applied in the MUVE-based

[1] The degree of authenticity can be varied depending on the design.

language teaching. It has been evident that MUVEs have a potential for more motivational learning (Tsai & Tsai, 2018), yet the effective pedagogical strategy needs to be investigated.

1.4. Outline of the thesis

The thesis is outlined as follows. Chapter 1 is a summary of the document, from which the readers can use it as a guideline to focus on certain parts. It introduces the context of the thesis within the avenues of literatures around sequence of guidance for effective learning, the aims and the significances of the thesis, the outline of the research, and the definitions of terms. Special attention on the literatures is the focus of chapter 2.

Chapter 2 provides a review of the relevant literature on productive failure pedagogical strategy applied in science, mathematics, and statistics. These studies reported that PF had helped students with procedural fluency, conceptual knowledge, and transfer. Reviews were also made on the digital game based EFL learning, that previous studies have not dealt with the pedagogical strategy in their experiments. Instead, they focused more on the internal factors of the game (affordances, modification) and external factors focusing on aspects like player versus non-player and frequencies of playing the game. This study proposes the use of Productive Failure (compared to Direct Instruction) as the pedagogical strategy of the game-like application in education, especially EFL learning.

Chapter 3 describes the methods of this study, including 1) participants, 2) setting of the experiment that includes real-world and virtual settings, 3) design of the experiments, 4) activity sequence of the treatment in the experiment (PF) and control (DI) groups, 5) data sources and analysis, and 6) materials in use.

Chapter 4 discusses the results of the study, which is organised based on the research questions. The first section of chapter 4 addresses the first research question, where it is shown that PF students outperformed DI students in procedural knowledge yet performed equally in declarative knowledge. The second section exhibits the answers to research question 2, describing the equal results of both PF and DI groups on the declarative and procedural knowledge of written and spoken description of places. The last section of chapter 4 shows the qualitative data of the research that were analysed based on the self-regulated learning framework.

Chapter 5 describes the findings of the research. The first section is the description of the findings presented based on the research questions. The first findings were related to the written assessment of describing places, while the second finding dealt with the spoken assessment on describing places. The last finding was about the qualitative finding on students' behaviours during the consolidation phase. The second section was the discussion

on the findings in relation to the previous studies on PF and SRL and the theoretical discussion based on PF and SRL theorising.

The last chapter elaborates the implication, limitation, and future research recommendation. The implication covered the area of classroom-based and online-based EFL. The limitation of the research was discussed in terms of the number of participants, the PF design, and the scopes of the SRL aspects being observed. The discussion on the limitation of the research is followed by recommendations on the future research either in foreign language setting or in the other subject setting.

1.5. Definition of Terms

The terms used in this dissertation are conceptualised into the definitions below.

1. Productive failure is a learning design comprising two phases, 1) idea generation and exploration and 2) consolidation knowledge assembly (Kapur, 2008; Kapur & Bielaczyc, 2012). The first phase allows students to solve a novel problem (communication tasks in this research) that targets a concept they have not yet learned, enabling prior knowledge activation and differentiation. The second phase lets students compare and contrast their solutions to the canonical (expert-like) solution, assisting the students to acknowledge their knowledge gaps.
2. Direct instruction (DI), or often known as explicit instruction, is an instructional strategy that prescribes teaching the critical concepts explicitly before problem-solving.
3. Declarative knowledge, often explicit, is knowledge of facts, concepts, procedures, and rules (Ullman, 2001). More specifically, declarative knowledge (in spoken and written description) in this research refers to the accuracy of vocabulary use, which is the relation between meaning and form (Paradis, 2009). Thus, knowledge that requires students to determine the English words from pictures or video chunks presented to them is classified into declarative knowledge. It applies to both written and spoken production of the target language. One example in written evaluation was like “A brown coloured _____ detailed with two blue vertical stripes in its centre, is stood upright against the stilt house.”, accompanied by a picture of surfboard. Meanwhile, in the spoken test was when students recalled vocabularies as they saw the representations in the virtual environment. These types of items ask students to recall vocabularies as a representation of things in the picture or the virtual world.
4. Procedural knowledge (used interchangeably with procedural fluency) is the ability to correctly apply the learned knowledge (Rittle-Johnson & Schneider, 2015). In this research, procedural knowledge refers to the accuracy of syntax (morpho-syntax, phonology, and

prosody) used in both written and spoken description of places (Paradis, 2009). Thus, knowledge that requires students to decide expressions beyond word-meaning representation is categorised into procedural knowledge both in the written and spoken production of the target language. These following items, "The light blue ocean calmly _____ (ripples) during the daytime." in written evaluation and "I can see beautiful red flowers living on top of a large rock." in spoken evaluation, measure students' procedural knowledge.

5. Self-regulated learning is an active constructive process whereby learners set learning goals and then attempt to monitor, regulate, and control their cognitive and metacognitive processes in the service of those goals (Azevedo, 2018). SRL strategies are translated into four phases: task definition, goal setting and planning, study tactics, and adaptation (Winne & Hadwin, 1998, 2008). This study focused on the study tactics as these are the observable behaviours during the consolidation (instruction) phase. Cognitive learning and metacognitive monitoring strategies are the exact strategies to analyse from the students' behaviour while accessing the instructional video.

6. Production Based Instruction (PBI) views production instead of comprehension of the target features as the source of acquisition (Shintani, 2015; Shintani, Li, & Ellis, 2013). Students' outputs become other learners' inputs. The key feature of PBI is corrective feedback to the students' production.

7. Inworld means "being connected to the Second Life servers and present in the Second Life world (also: online), anything that takes place within the virtual environment of Second Life. Used in this way it is spelled as one word with no hyphen, usage in line with email and website."

8. Stabilise, Simplify, Automatise, Reconstruct, and Complexify (SSARC) is a model of task sequencing proposed by Robinson (2010, 2015). This model suggests a simple-complex task sequencing after an explicit instruction to develop students' target language competency. Simple tasks help stabilize the newly acquired knowledge (lexical and grammatical forms); medium tasks give more opportunities for similar idea expression; and complex tasks pull students' potential for form-meaning connection in their spontaneous speech act (Allaw & McDonough, 2019; Lambert & Robinson, 2014; Robinson, 2006, 2015).

Chapter 2: Literature Review and Theoretical Framework

This chapter reviews the theoretical underpinnings of this research. It postulates two main approaches to learning new things in "game-like"-based English as a foreign language learning, specifically on degree and sequence of guidance. These approaches contradict each other not only on the need for guidance to foster students' learning but also on the right time to provide guidance. This research aims to compare the effectiveness of a combination of degree and sequence of guidance in foreign language learning.

This chapter starts with a review of existing literature addressing learning theories, especially on the degree and sequence of guidance to foster learning. Special discussion is devoted to PF and DI as two applications of a high and low combination of learning guidance. Task sequencing in language teaching as a similar application of sequencing guidance is also described. The following section discusses self-regulated learning as one of the success determinants in the learning process. The connection between self-regulated learning and both PF and DI is also elaborated.

The application of the two pedagogical strategies in MUVE-based English as a Foreign Language (EFL) learning is the next discussion. This research focused on the discussion of the help given to learning speaking in a foreign language, focusing on vocabulary and syntax development. The following section covers a discussion on self-regulated learning strategies in relation to the application of PF and DI pedagogical strategies. A proposal on how PF works in MUVE-based EFL activities is presented at the end of the chapter with the research questions and the hypotheses discussed.

2.1. Theories of Learning: Theoretical framework

Learning is a process of adaptation to the environment (Piaget, 2006) when there is a failure of fit between internal and external (environment) schemas. A process of conceptual change (Carey, 1999) marks the learning process in which accommodation is needed when there is a discrepancy between the students' concepts, which might be erroneous (Carey, 1999), messy, or ill-structured (diSessa, 2014), and the canonical concepts. The discrepancy will likely result in an impasse or temporary Failure, which is germane according to diSessa (2014b), and invoke mechanism or processes leading to more differentiated and complex structures. Thus, the primary purpose of pedagogical design is to create temporary Failure and to ensure the relevant guidance is provided to the students. The next two sections discuss guidance and PF as one of the pedagogical strategies that consciously design temporary Failure.

Temporary failures have been the primary students' experience in the "preparation for future learning" (Bransford & Schwartz, 1999), "impasse-driven learning" (VanLehn, 1988; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003), "inventing to prepare learning" (Schwartz & Martin, 2004), "assistance dilemma" (Koedinger & Aleven, 2007), and PF (Kapur, 2008). These instructional designs have a similarity that they let the students experience temporary Failure or an" impasse", which forces them to put all efforts into solving the learning problems. More learning efficacy may results after the Failure due to the learning strives (Loibl & Rummel, 2014a; Schwartz & Bransford, 1998; VanLehn, 1988; VanLehn et al., 2003). In line with the purpose of the study, PF is discussed after the discussion of degree and sequence of guidance.

2.2. Degree and Sequence of Guidance

The debate on the degree and sequence of instructional guidance in learning has been around for more than two decades. In terms of the degree of instructional guidance, two contesting ideas include low and high guidance of learning. In low guided learning, it is believed that effective learning happens when learners discover or construct essential information by themselves, not presented in an instruction (Kirschner et al., 2006). This way, they get ownership of the learning processes that lead to long-term retention and self-explanation ability (Jacobson et al., 2015). One example of low-guided learning is discovery learning, in which "the learner is not provided with the target information or conceptual knowledge and must find it independently and with only the provided materials" (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011, p. 2). In contrast, high guided learning claims that the concepts and procedures required for learning should be taught—especially for novices—instead of being left to the students to discover those procedures by themselves (Kirschner et al., 2006; Klahr & Nigam, 2004; Mayer, 2004). Learning is defined as a process of receiving information stored in the short-term memory (from the explicit instruction) and then sending it to the long-term memory (Mayer, 2004). The benefit of high-only guidance is a high degree of correctness and avoidance of initial Failure (Kirschner et al., 2006). An example of high-only guidance can be seen in tutored problems, worked examples, or erroneous examples (McLaren, Van Gog, Ganoe, Karabinos, & Yaron, 2016). In the area of language teaching, for example, students are presented with a model of a communicative act that they need to analyse the components with the help of the teacher. Guided practices then follow it until students are ready to do the task unassisted. In conclusion, in high-only guidance, the mastery of critical concepts and procedures is the key before solving problems or doing real-life tasks.

Despite their benefits, experts criticise their drawbacks in optimising learning. High-only guidance restricts the authenticity of the learning process, limits appropriate transfer, and decreases the ownership of the learning process (Jacobson et al., 2015; Kapur, 2008, 2016; Lee & Anderson, 2013). Within an explicit instruction, learning activities (tasks) are well-structured, which are far from the real-life condition. This condition restricts the authenticity of the learning process. When the process is not authentic, the transfer is hindered due to the far similarity of the classroom activities to the real-world conditions. Low-only guidance, in contrast, is criticised as a waste of time because the process takes time. Learning will unlikely happen as students' working memory become so occupied with problem-solving activities that they are not ready to encode new knowledge (Kirschner et al., 2006; Kirschner & Clark, 2004; Sweller, Kirschner, & Clark, 2007). Solving problems before the concepts and procedures are mastered will occupy the most working memory capacity, which should be used to encode new knowledge. Therefore, learning will not be likely to happen in this view. Alternatives to both extremes should be taken into account.

Jacobson et al. (2015) suggested a combination of low-high guidance and high-low guidance to optimise learning. In high-low guidance, explicit instruction or demonstration as an example of high guidance is presented to the students to acquire concepts and procedures. Practices are provided to ensure the concepts and procedures are mastered. A problem-solving activity is assigned when facts and procedures are fully acquired to minimise failures (Kirschner et al., 2006). For example, in an apprenticeship, students observe the expert or teacher doing a targeted task, which can be in the forms of modelling, coaching, and scaffolding (Gavriel, 2015). Additional structures like in a cognitive apprenticeship (Gavriel, 2015) are possible to ensure the acquisition of facts and procedures. Once concepts and procedures are mastered, problem-solving is assigned to the students (Jacobson et al., 2015).

The next combination is low-high guidance, which lets students do the ill-structured task independently at the beginning of the lesson, and then guidance is provided after. Cognisant to this sequence is Productive Failure, an instructional design allowing students to solve a novel problem prior to the subsequent instruction (Kapur, 2008; Kapur & Bielaczyc, 2012). This sequence of guidance is in line with Pea's (2004) statement that, in case guidance is needed, it should be given only to students who cannot do the task unassisted. Problem-solving prior to instruction ensures in which part is guidance needed to avoid the expertise reversal effect—less guidance is beneficial to more advanced learners (Kalyuga, Ayres, Chandler, & Sweller, 2003). More discussion on PF is covered in the next section.

This research takes the middling position, which is a combination of low and high guidance. The first reason is that combining the two extremes will likely maximise the benefits. For example, the low-high combination potentially decreases the learning inefficiency in the low guidance with the subsequent instruction – which is high-guided – following the low-guided activity like problem-solving. This condition also happens to high-low guidance. The low-guided activity will likely increase the authenticity of the learning process after the high guided activity, which has been criticised for the lack of authenticity. This study predicted that the low-high guided learning – like PF – will likely better optimise learning given that students have struggled before the explicit instruction.

2.3. Productive Failure: Theoretical Framework

Productive Failure is a learning design that allows students to solve a new problem that targets a concept they have not yet learned, followed by consolidation and knowledge assembly where they compare their solutions to the expert-like solution to learn the targeted concepts (Kapur, 2008; Kapur & Bielaczyc, 2012). The term "Productive Failure" was coined by (Kapur, 2008) in his seminal work in Cognition and Instruction journal on the same title. However, the idea of presenting Failure in learning has been evidenced in the impasse-driven learning (VanLehn, 1988), preparation for future learning (Schwartz & Bransford, 1998), inventing to prepare for future learning (Schwartz & Martin, 2004), and model-eliciting activities (Doerr & Lesh, 2003). These models suggest that learning does not happen until students experience Failure. In impasse-driven learning, for example, students interpret and try solving a problem without helps. While solving the problem, they will surely get stuck at some point. To overcome this stuck problem, they will try some repairs until they get out of the problem. This repairing activity constitutes real learning (VanLehn, 1988). The process is similar to PF, in which interests have been flourishing.

A number of studies on PF have been conducted in many areas, but language teaching. Those areas include physics (Jacobson et al., 2020), mathematics (e.g., Kapur, 2011; Loehr et al., 2014), statistics (Newman & DeCaro, 2019), and social sciences (e.g., Holmes, Day, Park, Bonn, & Roll, 2014; Jacobson et al., 2017; Pathak, Kim, Jacobson, & Zhang, 2011). The majority of the findings revealed the superiority of PF over DI on the procedural knowledge (e.g., Kapur, 2015), conceptual knowledge (e.g., Kapur, 2014a; Lai et al., 2017), and transfer (e.g., Jacobson et al., 2017; Kapur, 2014a). These studies prescribed the problem solving prior to instruction, which is the characteristic of PF pedagogical design.

There are two phases in the PF design, (a) an initial generation *and exploration* phase and (b) *consolidation and knowledge assembly* phase (Kapur, 2014a; Kapur & Bielaczyc,

2012). The first phase lets students solve a set of novel problems, which should lay within the students' *zone of proximal development* (Jacobson et al., 2015) to avoid frustration (Kapur, 2014b, 2016). As students are not experts, they have limited conceptual or knowledge resources to solve new problems. In the condition that external help is not available, they can only rely on their prior knowledge, in which Failure is highly likely to happen. This Failure is then treated in the second phase, the *consolidation and knowledge assembly.* In the consolidation phase, comparing and contrasting students' representation and solution methods (RSMs) to the canonical solution facilitates knowledge gap recognition (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), while the knowledge assembly helps with the schema abstraction (Jacobson et al., 2020). Practically, students will likely join the explicit instruction on the critical concepts or procedures by bringing misconceptions which resulted from the problem-solving activity. They compare their incorrect concepts to the correct ones during the instruction. The revised concepts are then stored in abstraction in the long-term memory, which are available for retrieval to transfer in different contexts.

To successfully apply PF in the class, the initial problem-solving tasks should provide opportunities for the students to explore possible strategies to generate multiple solutions. Therefore, (a) the tasks should make prior knowledge activation possible (Kapur, 2014b, 2016); (b) students must themselves generate and explore solutions rather than simply be presented with peer's solutions (Kapur, 2014b); (c) the tasks should be challenging enough but not so challenging that makes the students give up; and (d) the teacher should point out the critical features of the targeted concepts to compare and contrast students' generated solutions to the canonical solutions (Kapur, 2016).

The critical mechanisms in productive Failure include prior knowledge activation and differentiation (Kapur, 2014b, 2015, 2016; Kapur & Bielaczyc, 2012), knowledge gap recognition (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), and schema abstraction (Jacobson et al., 2020). First, prior knowledge activation and differentiation is facilitated in the first and second phases of PF design, the idea generation and exploration. Prior knowledge activation is promoted during the idea generation and exploration provided students cannot access external supports when solving a novel problem before instruction (Kapur, 2015; Kapur & Bielaczyc, 2012). As students are not taught the main concepts of the targeted topic yet, they might experience a failure or an "impasse,"—the condition when students get stuck in solving problems (VanLehn, 1988; VanLehn et al., 2003). Learners will seek for their prior knowledge, the only sources they have, to solve the

initial problem. Once their prior knowledge is activated, the extent to which they will learn from the instruction can be predicted (Kapur & Bielaczyc, 2012). Additionally, prior knowledge differentiation is supported by the comparing and contrasting activity in the consolidation—among students' solutions or between students' solutions and the canonical solution (Kapur, 2015). This process invokes learning in that students can acknowledge the critical features of the concepts. In conclusion, the first mechanism of PF facilitates prior knowledge activation to help students' learning, which will then support the second mechanism.

Second, knowledge gap recognition happens as soon as prior knowledge has been activated from the first activity. A general knowledge gap recognition is facilitated in the first phase when students compare their solution to their friends' while specific knowledge gaps recognition is promoted in the second part of the consolidation phase—when students' solution is compared to the canonical (expert's) solution (Loibl & Rummel, 2014a, 2015). The comparison between students' solution to the canonical solutions embellishes the characteristic of instruction in PF. Specific knowledge gaps recognition has been evidenced to facilitate learning as students can focus on subsequent instruction (Chi, 2000; VanLehn, 1988). This mechanism invokes better learning, as students are ready for future learning considering schema abstraction is facilitated.

Last, schema abstraction connects the new knowledge with the old one stored in the long-term memory (Jacobson et al., 2020; Kapur & Bielaczyc, 2012). Prior knowledge is activated and differentiated in the first phase of PF (Kapur, 2014b, 2015, 2016; Kapur & Bielaczyc, 2012), within students' old schema. The activity of comparing and contrasting in the consolidation phase enables knowledge gap recognition that shows the position of the old schema relative to the targeted new one. These knowledge gaps ease the transformation of the old schema to the new one provided that students are prepared for the subsequent instruction (Jacobson et al., 2020). DI students, in contrast, might not activate their prior knowledge as they jump to the instruction. They more likely experience "false mastery," that they feel they have mastered the content of the instruction. False mastery experience will lead to missing attention to critical features in the instruction, which might further result in less facilitated schema abstraction.

This research employed PF design (Kapur, 2008; Kapur & Bielaczyc, 2012) comprising the timing and the design of instruction. In the idea generation and exploration, a communicative task of describing an outdoor place was assigned. This communicative task is slightly above the level of the students' skills. Students were expected to activate their prior

knowledge and experience failure given they did not understand some vocabularies, syntax, and prepositions. The consolidation and knowledge assembly presented a canonical description as the model at the first section and a discussion on the critical features of describing places after that. The model functions as the standard for students to compare and contrast their descriptions in idea generation and exploration. The description of the critical features supports students' schema abstraction by filling in the knowledge gaps with the features presented in the instruction. This is called schema abstraction.

At the beginning of the instructional video, students were presented with a canonical description presented by an L1 English speaker. This allowed students to compare and contrast their description and the canonical one. At this stage, students could have put their description and the canonical description side by side in their mind so that they could compare. The results of the comparison could help them acknowledge their knowledge gap, which can be an initial condition for deeper learning during the instruction. The instruction discussed the critical features of describing places, especially the vocabularies and syntax, which were the primary investigation of this research. It is expected that the students paid more attention to the features to fill their knowledge gaps. As they attended the instruction to fill their knowledge gaps, the meaningful learning experience could take place. This activity is believed to support schema abstraction process (Jacobson et al., 2020). In the following sections, we will see what studies have been carried out on PF.

2.4. Productive Failure: Literature Reviews

Previous studies on PF have reported its superiority over DI on students' procedural fluency, conceptual knowledge, transfer, and mental efforts, carried out in mathematics (Kapur, 2011; Kapur & Bielaczyc, 2012), statistics (Kapur, 2014a, 2015; Schwartz & Martin, 2004), and science (Holmes et al., 2014; Jacobson et al., 2017; Kapur, 2008; Kapur & Kinzer, 2009; Kennedy-Clark et al., 2009; Pathak et al., 2011). The next paragraphs discuss the main findings of PF studies.

2.4.1. Productive Failure and procedural Knowledge

Several studies have revealed the effect of PF on procedural fluency with varied results (Kapur, 2012, 2015; Loibl & Rummel, 2014a). For example, in Kapur's (2015) study, 133 ninth-grade students learned Standard Deviation in a quasi-experimental design. Students in the PF group solved a complex data analysis problem in triads without guidance. In the consolidation phase, the teacher compared students' solution and then modelled and worked through the canonical solutions. This is followed by them solving three data analysis problems, and the teacher discussed the answers with the class. In contrast, in DI, the teacher

explained the concepts by using two sets of worked examples followed by problem-solving activities. They solved an isomorphic problem and discussed the solution in the class. Students were asked to solve three isomorphic data analysis problems as practised in the second phase and discussed it with the teacher. In the third phase, they solved three problems in triads and discussed the canonical solution with the teacher. The results showed a significant difference in procedural fluency and conceptual knowledge with the larger effect size on conceptual knowledge.

Loibl and Rummel (2014a, 2014b) reported the effect of PF on procedural fluency differently. Fluency was measured by the time needed to solve a problem. In Loibl and Rummel (2014a), for example, 98 tenth graders students were assigned into I-PS, $I_{contrast}$-PS, and PS-$I_{contrast}$ in a real school setting. They learned the concept of variance based on their groups. I-PS got their instruction before solving a problem, $I_{contrast}$-PS got their instruction with comparison and contrast, while the PS-$I_{contrast}$ let the students solve a novel problem before instruction, which involves comparison and contrast. The results indicated that I-PS scored higher on procedural skills. They argued that procedural skills improve due to the number of practices, not the order of instruction. This result was the same compared to both guided or unguided problem solving (Loibl & Rummel, 2014b).

While both studies claimed that problem-solving before instruction was beneficial for procedural fluency, they both had a slightly different design. In Kapur & Bielaczyc (2012), the PS-I was compared to I-PS while in Loibl & Rummel (2014b), the PS-I was compared to PS-I with or without comparison and contrast. While Loibl & Rummel (2014b) suggested that procedural skills were affected by the number of problems solving not the sequence of activities, this research takes the position that sequence does affect the procedural knowledge. Provided the consolidation and knowledge assembly in the PF helps the schema abstraction better (Jacobson et al., 2020), procedural fluency will likely be affected. The fluency of using the new knowledge in a new situation is supported with the ease of recalling the schema from our long-term memory, which is helped better when the schema is abstracted more straightforward. Therefore, this research used PF's original design (Kapur & Bielaczyc, 2012) in that both the number and the type of activities are the same with different sequences, PS-I vs I-PS. A more explicit description of the design is presented in chapter 3.

While Loibl & Rummel (2014b) measures procedural fluency as the duration of task completion and Kapur measured procedural skill with how fluent a student computes and interprets the canonical solution (Kapur, 2012), this research defined procedural skill based on Paradis (2009). Procedural skills refer to Paradis' (2009) work, where vocabulary

encompasses the relation between meaning and form (both written and oral) and is declarative, while syntax (morpho-syntax, phonology, and prosody) is of procedural knowledge.

2.4.2. Productive Failure and conceptual knowledge

PF-related studies found PF's superiority over DI to improve students' conceptual knowledge (e.g., Chowrira, Smith, Dubois, & Roll, 2019; Kapur, 2012, 2014a, 2014b, 2015; Kapur & Bielaczyc, 2012; Lai, Portolese, & Jacobson, 2017; Loibl & Rummel, 2014b; Newman & DeCaro, 2019) facilitated by three mechanisms. In Lai et al.'s (2017) study comparing low-high (LH) to high-low (HL) pedagogical structures, two groups of engineering students were trained to solve a problem. They were randomly assigned to LH and HL groups. Students did an authentic task and a course unit on statistical process control in a different order. The LH group solved the problem before participating in a lecture while the HL group received a lecture before solving the problem. A task on conceptual knowledge was assigned in the pre-and post-test while a transfer task was assigned on the post-test only. LH group outperformed HL group both in conceptual knowledge and transfer (Lai et al., 2017). The results support the previous studies on PF that it helps students improve conceptual knowledge, an important goal of learning.

2.4.3. Productive failure and transfer

Significantly better transfer, the main characteristic of meaningful learning (Mayer, 2002), has also become the main finding of PF studies. The results of PF-related studies on transfer revealed the superiority of PF in far transfer—between domain transfer (Jacobson et al., 2017; Kapur & Kinzer, 2009), and near transfer—within domain transfer (Kapur, 2014b; Lai, Jacobson, & Goldwater, 2018; Lai et al., 2017). Solving real-world problems in the idea generation and exploration phase might invoke transfer skill due to the authenticity of the learning process, where students have experienced a complex ill-structured problem (Jacobson et al., 2017). PF students' ability to transfer is facilitated as the learning events share a significant number of similar stimuli with the transfer event or the real use of the knowledge (Royer, 1979) provided they solve ill-structured problems before or after the instruction. Thus, PF students have higher opportunity to deal with similar structure with the real use of the knowledge or skill in the first phase.

2.5. Direct Instruction: Theoretical Framework

Direct instruction (DI), or explicit instruction, is an instructional strategy that includes explicit teaching of the critical concepts before solving real-life problems. In DI classes, teachers tell, show, model, demonstrate, and teach the vital concepts and skills to be learned

(Baumann, 1988) based on a solid plan (Anderson et al., 1988; Engelmann, 2007). Learning is defined by reducing errors and misconceptions among students so that there is no room for failures (Kirschner et al., 2006; Klahr & Nigam, 2004). The mastery of critical concepts from the instruction is believed to ease the cognitive resources and support accurate knowledge and procedure acquisition (Lee & Anderson, 2013). Merely solving problems for learning or prior to learning may hinder knowledge acquisition as working memory is not available for learning (Kirschner et al., 2006). Teaching students the new knowledge and procedure will help them solve the problem correctly. Their working memory becomes free of disengagement and frustration arising from problem-solving before instruction (Hardiman, Pollatsek, & Well, 1986). The DI pedagogical strategy shares the same features with the weak-form of *Task-Based Language Teaching* (TBLT) application.

Task-based language teaching (TBLT) is an example of DI applied in language learning, in which the mastery of concepts is assured in the instruction prior to real-life language tasks. Meanwhile, the strong form application of TBLT (Ellis, 2019) is an example of low only guidance. The majority of task sequencing approaches in the weak-form application of TBLT, e.g., Robinson's (2010) and Robinson and Gilabert's (2007) SSARC or Triadic componential Framework, follows the high-low guidance principles. Students were asked to do tasks either in simple-complex or complex-simple sequences after explicit instruction, which generally presents the vocabulary, syntax, and pragmatics issues needed for task completion. Therefore, the weak-form TBLT application corroborates DI regardless of the task sequences.

2.6. Instructional Sequence in Language learning: A literature Review

This research was informed by a weak form of TBLT application, in which explicit instruction is still needed despite the communicative tasks. A weak form of TBLT—usually called task-supported language learning (Ellis, 2019)—sees tasks as a vital part of language instruction embedded in more complex pedagogic contexts. They are essential but might be preceded or followed by focused instruction on grammar, vocabulary, or pronunciation provided it is complementary to task completion. This weak form's benefits are on the accuracy and durability of the acquired knowledge (Lee & Huang, 2008). The accuracy supported as forms are introduced explicitly and practised in guided activities prior to the production. Durability is enforced from a number of guided practices on the specific forms and executed in the free activity in the production. In conclusion, explicit instruction is the main characteristic of the weak-form application of TBLT; in which the timing depends on the task sequence.

Studies on task sequencing in language teaching were first marked by a number of studies on task sequencing in Baralt, Gilabert, and Robinson's (2014) volume with contrasting results. On one side, performing tasks in a complex-simple sequence was evidenced to improve language-related episodes (LRE) (Baralt, 2014) related to target language development. On the other side, studies found that simple-complex task sequence improved language development better than complex-simple task sequence (Allaw & McDonough, 2019). However, unlike those two studies, Malicka (2014) and Lambert & Robinson (2014) found no effect of task sequencing on students' language development. The next three paragraphs describe the primary studies on task sequencing in TBLT one by one.

In Baralt (2014), 94 Spanish learners were grouped into complex-complex-simple (CCS), simple-simple-complex (SSC), simple-complex-simple (SCS), and complex-simple-complex (CSC) task sequences in both online and offline contexts. A story retelling tasks were sequenced based on the complexity; read a story in the first language (simple)—retell in the target (scaffold: 12 cards—six with a brief blurb story in L1, six with comic strip) (complex). Instruction on subjunctive was given before students carried out three tasks based on the sequences. A written and oral pre-test was assigned before the treatment, and a post-test and a delayed post-test were assigned after the treatment. The tests were written and oral story retelling in L1 individually and then in L2 collaboratively. The accomplishment of the task to cover all aspects determined the criteria for successfully finish the tasks. Results showed that complex-simple task sequencing helps students improve language learning episodes (LREs) and language development in an offline class. In the online context, no differences were found between simple-complex and complex-simple task sequences.

In Malicka's (2014), 50 English as a second language learners were assigned in six different groups, with 25 learners in simple-medium-complex task sequence while 5 students each were assigned in one of the sequences (simple-+complex-complex; complex-simple-+complex; complex-+complex-simple; +complex-complex-simple; +complex-simple-complex). The complexity was determined by the number of elements (characteristics of clients and rooms) and reasoning demands (describing, recommending, apologising, and justifying). Performance measurement on room recommendations, which assessed complexity, accuracy, and fluency (CAF), was taken from the simple, medium, and complex tasks. The results confirmed the previous study that task complexity affected CAF, but not structural complexity, and revealed no difference in the language development between simple-complex and random task sequences.

Informed by Robinson's (2010) SSARC model (Stabilise, Simplify, Automatise, Reconstruct, and Complexify), Allaw and McDonough (2019) assigned 42 French students into simple-complex and complex-simple sequencing groups. The complexity was determined based on the Triadic Componential Framework (Robinson & Gilabert, 2007) by manipulating the resource-dispersing (± task structure) and resource directing (± spatial reasoning) dimensions. In simple-complex task sequence, instructional guidance was provided in paragraph format and explicit instruction on how to write a paragraph for the simple task, which was usually called explicit instruction. For the medium task, students had to describe a park's image without any instructions about paragraph format or guidelines. The complex task eliminated the help in the first simple and medium tasks leaving a primary task of describing places relying on the picture in students' mind. The results revealed that simple-complex task sequencing helped students learn vocabulary and grammar better and maintained performance over time. Unlike the simple-complex sequence, students assigned to the complex-simple task sequencing improved the vocabulary and grammar only in the immediate post-test.

There are two interesting notes to consider from the sample studies on task-sequencing in TBLT applied in a foreign language. First, the number of students in the studies were either not balance between groups like in Baralt's (2014) or too small to draw a robust conclusion like in Allaw & McDonough's (2019) and Malicka (2014, 2020). Second, the comparison between the experimental and the control groups was not equal. For example, explicit instruction was provided prior to the sequenced tasks in the experimental condition (simple-complex sequence) but not in the control condition (complex-simple sequence) (Allaw & McDonough, 2019). Another example is the comparison between simple-complex task sequences and a single complexity (simple only, medium only, or complex only) (Malicka, 2014). These designs may not be useful to compare the effectiveness of task sequences due to imbalanced structure. The results may not be the effect of the sequence, but other aspects like the explicit instruction or task complexity. A fair comparison of task sequences should include the same activities and tasks yet sequenced differently. Figure 1 might be a fair alternative design to compare task sequencing studies.

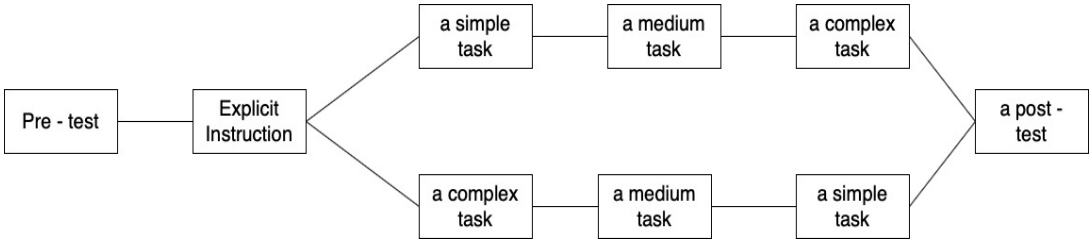


Figure 1 An alternative comparison design for task sequencing studies

This alternative design is similar to the comparison between PF and DI pedagogical strategies. While complex-simple task sequencing is identical to low-high guidance, there are essential differences in the instruction and complex tasks. The problem-solving in the low – high guidance (e.g., Productive Failure) lets students solve a complex novel problem without helps with the purpose of prior knowledge activation and differentiation (Jacobson et al., 2020; Kapur, 2012; Loibl & Rummel, 2014a). In contrast, performing a complex task in task sequencing was not for prior knowledge activation provided that explicit instruction is already available before the task (e.g., Allaw & McDonough, 2019). In the PF explicit instruction, comparison and contrast between students' solution and the canonical solution become the essential procedure (Kapur & Bielaczyc, 2012; Sinha & Kapur, 2019), which facilitates specific knowledge gap recognition. The instruction in TBLT, in contrast, has no particular procedures, regardless of the sequence. A comparison and contrast between students' task performance and a model task does not occur in the instruction provided that the tasks are carried out after the instruction. The next paragraph describes the design of this research.

This research applied the PF pedagogical strategy in foreign language learning with the activity sequence of complex task 1 – instruction – complex task 2. The pedagogical strategy has been applied successfully in science and mathematics (Kapur, 2014a; Loibl & Leuders, 2019; e.g. Loibl & Rummel, 2014a). PF's design (Kapur & Bielaczyc, 2012) was proposed to be an alternative to the pedagogical strategy commonly used in EFL – the simple-complex or complex-simple task sequences in TBLT. Both sequences are preceded by explicit instruction, which discusses the concepts and procedures to carry out the tasks. The explicit instruction prior to task completion has been criticised for restricting students' creativity because students tend to rely too much on the model to finish the task (Ellis, 2018).

Table 1 compares the PF and TBLT design, in which the main difference is on the design of the instruction and the guidance. PF prescribes comparing and contrasting students' solution (task performance) to the canonical solution (canonical task performance) in the instruction phase, while TBLT does not specify the activity during the instruction. However,

an inductive approach is suggested to help students with vocabulary and grammar in the pre-task activity (Li, Ellis, & Zhu, 2016) to get ready to perform the main task. Similar to DI, TBLT puts instruction at the pre-task cycle to prepare students in the primary task cycle.

Table 1 The comparison of Task-based Language Learning and Productive Failure

Aspects	Productive Failure	TBLT
Activity Sequence	A real-life task – instruction – a real-life task	Simple task – medium-simple tasks – more complex (real-life) tasks
Instruction design	Compare and contrast between students' solutions (failed and suboptimal) to the canonical solution followed by direct/explicit instruction	Direct/ explicit instruction – delivering the content of the materials
Guidance	Low – high	High – low (little was carried out in low – high
Phases	Idea generation and exploration Consolidation and knowledge assembly	Priming (pre-task) Primary task Follow up

2.6. Grammar Instruction: Theoretical Framework

This project used Bock & Levelt's (1994) theory of language production, which consists of two processes – grammatical encoding and phonological encoding. Grammatical encoding comprises selecting appropriate lexical concepts (entries in the speaker's vocabulary) and the assembly of a syntactic property (Bock & Levelt, 1994; Dell, Chang, & Griffin, 1999). Appropriate lemmas (basic vocabularies) are selected and then assigned their functions in the sentence. Different assigned functions allow different syntax (e.g., SV-agreement, plural-singular, noun phrase). The phonological encoding used the encoded vocabulary and syntax in the production of well-spoken sentences. This research limits the work on comparing two pedagogical strategies on EFL learner's grammatical encoding ability.

Grammatical ability is defined as the ability to use grammar as a communication resource to create spoken and written texts (Richards & Reppen, 2014). Thus, measuring grammatical ability is measuring production skill. In this research, the grammatical ability includes lexical and syntactic accuracy in both the written and oral descriptions of places resulting from taking part in either PF or DI pedagogical strategies. The vocabulary was considered declarative, while the syntax is procedural (Paradis, 2009). Therefore, vocabulary is declarative knowledge provided it is assessed simply based on a meaning association

between an object and the lemma (basic vocabulary). Unlike vocabulary, the syntax is procedural as it deals with the meaning association and the grammatical representation due to the function assigned to the vocabulary in a clause. For example, in the test item:

	"A brown coloured _____ (6), detailed with two blue vertical stripes in its centre, _____ (7) upright against the stilt house," there are declarative and procedural questions. Answers: 6 a surfboard 7 is stood

Item number 6 assesses declarative knowledge as students only need to recall the word "surfboard" from a visual stimulus in the picture. Number 7, on the other hand, measures procedural knowledge as students need to recall a lemma "stand" functioned as the predicate in the passive form. In order for students to own grammatical knowledge and grammatical ability, effective instruction is needed.

2.6.1. Grammar instruction: aA literature review

While the purpose of EFL learning is developing the four macro skills (listening, reading, speaking, and writing), the sub-skills needed for the development of macro-skills need considering as well (for complete sub-skills, see Tarone, 2005). Grammar (which includes vocabulary and syntax) has been the challenge for foreign language learners, including Indonesia (Cahyono, 2008; Kim, 2020). Without vocabulary, nothing can be communicated (Wilkins, 1972), and without syntax, communication can be less effective (Bock, 1986). Therefore, effective pedagogies for grammar development are essential to search for.

This research used explicit instruction as a method of supporting vocabulary learning to match with the PF design. While learning L1 vocabulary and syntax is effortless (implicit learning), it is not for L2, which might need explicit learning (Mar-molinero & Stevenson, 2006). The implicit (or incidental) learning enthusiasts (for example, Reynolds, 2015; Wang, 2013) believe that vocabulary and grammar were acquired unconsciously by getting involved in communicative acts of reading or watching. The explicit (or intentional) learning, on the other hand, suggests that effective vocabulary and grammar acquisition is a result of conscious learning (Schmitt, 2008; Tammenga-Helmantel, Arends, & Canrinus, 2014). While implicit learning has the advantages of more meaningful learning, it has been criticised for the indefinite target vocabulary and its lengthy-time (Ellis, 2018). Similarly, explicit instruction has its advantages of efficiency and accuracy yet been criticised of the retention. Have studies so far take these issues into account?

This research employed the keyword method (Sagarra & Alba, 2006) to introduce vocabularies in the consolidation. The vocabularies were presented to students in both groups by showing pictures with spelling annotation while mentioning the vocabularies in contextual sentences. Students were expected to associate the pronunciation, spelling, and meaning through the visual representation of the target vocabularies. A current review of studies on vocabulary instruction reported that L1-L2 pair presentation with or without retrieval practice (e.g., Barcroft, 2007, 2009) had dominated the laboratory-based treatments in vocabulary learning (Rice & Tokowicz, 2020). The review suggests that repetition improves form acquisition and recall while keyword method and learning in context sentences helps with L2 meaning acquisition and recall (Rice & Tokowicz, 2020). For example, a study investigating three vocabulary interventions suggests that the keyword method was superior to repetition and semantic mapping vocabulary retention (Sagarra & Alba, 2006). Comparing rote memorisation, semantic mapping, and keyword method, Sagara and Alba (2006) randomly assigned 778 beginning *SL* learners into the three groups. They found that deeper processing via form and meaning association (i.e., keyword method) resulted in better retention than rote memorisation and multiple-meaning associations.

In line with the studies on classroom-based grammar instruction, a myriad of studies was also carried on the online-based intervention. The next section is the discussion on the investigation of the multi-user virtual environments-based grammar instruction.

2.6.2. Grammar instruction in multi-user virtual environments (MUVE)

From the establishment of *SL*, several studies have been carried out on the use of *SL* in foreign language learning (Chen, 2010; Cooke-Plagwitz, 2008; Godwin-Jones, 2011; Huang et al., 2012; Lan et al., 2013; Peterson, 2011; Rahayu & Jacobson, 2012; Wang, Petrina, & Feng, 2017; Wehner, Gump, & Downey, 2011; Zheng et al., 2009). The common themes of previous studies were SL's affordances, SL's potentials in improving target language performance, the teachers' and students' perspectives, and the challenges of integrating *SL* in foreign language learning.

Previous studies suggest that Learning in *SL* increased immersion and active participation (Liou, 2012). Users can interact with a variety of norms of social interaction (Steinkuehler, 2006), and they can have the opportunity to experience life-like social interaction (Cooke-Plagwitz, 2008). Also, learning in *SL* supports authentic communication experience (Lan et al., 2013; Rahayu & Jacobson, 2012), which are a potential success factor of foreign language learning. Investigated in a case study, participants in (Peterson, 2011) undertook tasks successfully via the target language interaction in Active Worlds. The

learners' interaction was supported by the orchestra of task type, socio-linguistic factors, context, and technical affordances.

Learning in MUVEs enhances students' affective factors (e.g., self-efficacy, motivation, and anonymity) that may support oral development in foreign language learning. Studies suggest that MUVEs improve self-efficacy (Rahayu & Jacobson, 2012; Zheng et al., 2009) and reduces anxiety (Wehner et al., 2011) which then enhances motivation to learning foreign languages (De Lucia, Francese, Passero, & Tortora, 2009; Wang et al., 2017; Wehner et al., 2011). A research project conducted by Rahayu & Jacobson (2012) involved *SL* for an EFL activity to enhance foreign language speaking. This MUVE was intended to support an authentic use of the target language for four-class sessions and three to four extra-class sessions, and it was found that students increased their speaking self-efficacy. For example, one of the participants helped a friend to change appearance in *SL*. He described how to buy, take off, and alter the outfit in both voice and text messages. In the interview, he commented that he felt confident that his spoken English was understood world-wide, which suggested he had speaking self-efficacy (Bandura, 1997). He also felt the genuine need to help his friend by using any expression he needed to convey his intended meaning (for a complete description, see Rahayu & Jacobson (2012)).

The next affective factor supported by MUVE-based class attendance was motivation. In a 10-session Spanish course investigating students' motivation to learn Spanish, students carried out a project that let them interact with Spanish native speakers via text and voice chats compared to those working in a traditional class (Wehner et al., 2011). A motivation survey after the treatment was given to both groups. The result shows that students improved their motivation to learn Spanish due to reduced anxiety during language practice. Anonymous feelings facilitate anxiety reduction due to avatar representation, text chat, and unique names (Dickey, 2005, 2011). A participant in Rahayu & Jacobson's (2012), for example, confidently spoke to avatars in *SL* while he was joining a four-week MUVE-based English course. He was usually shy in the classroom-based English lessons. Once the researcher asked him to talk in front of the class, he could express his ideas. Not only that, but he also laughed for a long time until he was asked to sit back to his chair.

The supports of MUVEs on students' affective factors might be one reason for their language development. In EFL speaking, for example, taking an online class in MUVE increases students' oral performance (Jauregi, Canto, de Graaff, Koenraad, & Moonen, 2011; Lan, 2014; Lan et al., 2013; Rahayu & Jacobson, 2012). The improvement was evidenced both from students' subjective ratings (Lan, 2014; Lan et al., 2013) and from an objective

measurement of the complexity of utterances and length of oral production (Rahayu & Jacobson, 2012). The improvement is related to the improved self-efficacy, motivation, and immersive experience due to the sense of presence, which is an essential factor in successful EFL learning (Kontogeorgiou, Bellou, & Mikropoulos, 2008; Mikropoulos, 2006). Presence enhances "first-person psychological activity occurring when people interact directly with worlds, whether real or virtual" (Winn, 1993).

In another MUVE-based EFL research, (Wang, 2015) investigates task-based language teaching courses in *SL*. In delivering the online course, the teacher uses the sequence of pre-task, during-task, and post-task. In the pre-task, learners are taught the targeted language expressions to use in the during-task phase, which helps the learners' fluency and accuracy while doing the task (Bygate, 1999). The study focuses on the teacher's role in motivating the students' participation in SL's online course. The findings suggest that during the pre-task, the teacher introduces types and tokens of discourse functions and gives the technical and social role. During the task, she focuses on motivating students to participate, monitoring student activities, and providing task supports. During the post-task, the teacher functions herself as a language guide. Despite the mention of task-based language teaching, it uses PPP sequence in the design, which is not relevant to the principles of task-based language teaching.

Specific to grammar instruction, digital game-based vocabulary learning, in which explicit instruction is suggested (Khezrlou & Ellis, 2017), has been reported to significantly enhance vocabulary compared to traditional instructions (Tsai & Tsai, 2018). This meta-analysis suggests that there was no evidence supporting Chiu's (2013) finding that Computer-assisted Language Learning (CALL) without games affects more compared to CALL with games for vocabulary learning. Since a significant effect of digital game-based learning (DGBL) was found in Chiu, Kao, and Reynolds (2012), the negative effect might come from the other factors other than the game itself.

In a meta-analysis of game-based vocabulary learning, investigations have been done on the game-internal factors (e.g., task vs drill-based games; base-game vs enhanced game) and game-external factors (e.g., player vs non-player, out of class vs in class, high vs low gaming frequency) (Tsai & Tsai, 2018) that support vocabulary learning. In game-internal factors, findings suggest that a task-oriented game was more powerful to support vocabulary learning as it stimulates critical thinking and problem-solving (Baralt & Gomez, 2017; Chen, Tseng, & Hsiao, 2018). However, games with added feature support better when the setting is more informal, and the assessment is productive (Tsai & Tsai, 2018). In the external factors,

the players' educational and proficiency levels have been the significant moderating factors of the DGBL success to support vocabulary learning. Higher L2 proficient university students gain higher vocabulary scores (Abraham, 2008; Pulido, 2003). Appropriate prior knowledge was the reason for the improvement as it supports the connection between words which then accelerates vocabulary gains.

The majority of DGBL-based studies compared it to traditional vocabulary learning, which resulted in a high effect size (Tsai & Tsai, 2018). In this design, all factors can influence the success of DGBL; game-design factors, game-internal factors (genre, topic, characters, contextual info, sound, music, graphics, rules, which potentially affect students motivation, engagement, immersion, and learning (Cairns, Cox, & Nordin, 2014). The second design compares base game to enhanced game (a game with added features like scaffolding). This design looked at internal game factors (mostly the added augmented features). Studies in this group resulted in medium effect size, showing that enhanced games gave more benefits to vocabulary learning. The third design can be a digital game compared to conventional (classroom-based game). The focus of the investigation was the game interface; how it affects vocabulary learning. Results showed medium to a high effect size of the studies in this group, indicating that digital gaming was more beneficial to vocabulary learning than the traditional games. The last design looked at factors outside the game as it employed the same games. Studies on factors outside the game to date compared between players and non-player, out of class and in class, high and low gaming frequency (Mohsen, 2016; DeHaan, Reed, & Kuwada, 2010; Sundqvist & Wikström, 2015). The findings suggest that frequency plays a vital role in the effectiveness of DGBL for vocabulary as well. DGBL cannot be used too long, despite its motivating factor, as it can create fatigue and boredom (Segers & Verhoeven, 2003) and short-term memory loss (Chiu, 2013a; Cowan & AuBuchon, 2008), which is counter-productive for vocabulary learning.

This research is intended to fill the gap of the limited study conducted on the pedagogical strategies used in DGBL. There are two main options for research that would explore this issue of identifying effective pedagogies for EFL involving MUVE systems: (a) employ existing EFL pedagogical approaches, and (b) employ innovative pedagogical approaches used in other subject areas. For the first option, the TBLT-based technique (Cook, 2008) in both the task sequences could be implemented in a MUVE system while for the second option, a recent learning design PF (Kapur & Bielaczyc, 2012) can be tried out. The reason is that PF has received considerable research attention for learning in scientific and mathematical knowledge (Jacobson et al., 2020, 2017; Kapur, 2008, 2014b, 2016; Sinha &

Kapur, 2019). This approach has also been successfully used in a virtual learning environment (similar to a MUVE) for helping students to learn about scientific inquiry (Jacobson et al., 2015).

2.7. Self-regulated Learning (SRL)

The working definition of self-regulated learning in this research is a constructive and multidimensional process involving cognition, metacognition, motivation, and affect to enhance academic achievement based on the previously set learning goals. This definition is summarised from Azevedo, Behnagh, Harley, and Trevors's (2010), Dörnyei's (2005) and Pintrich's (2000) definitions. Within self-regulated learners, there are common characteristics that they are active and efficiently manage their learning through monitoring and strategy use toward achieving their learning goals (Pintrich, 2000; Winne, 2015; Winne & Perry, 2005). Motivation, monitoring, and learning strategies were the SRL focus on this project.

This research views SRL as a dynamic and context-specific process rather than a static attribute. SRL is context-dependent, and it may vary across and within the learning activities (Bråten & Samuelstuen, 2007; McCardle & Hadwin, 2015) and even fluctuates over the course of learning (Moos & Azevedo, 2008a). In Bråten and Samuelstuen, for example, students enacted difference strategies for an exam in contrast to reading for a class assignment (Bråten & Samuelstuen, 2007). In language learning, for example, students might use different learning strategies for vocabulary mastery compared to pronunciation fluency. This leads to the choice of "online" observation as the data-gathering technique. It is assumed that a self-reported questionnaire will not capture such context-specific behaviours as it measures SRL as a static trait (Rovers, Clarebout, Savelberg, de Bruin, & van Merriënboer, 2019) before or after the learning journey.

A number of experts have proposed macro-level models of SRL. Pintrich (2000) suggested three phases of self-regulated learning: namely (1) task identification and planning, (2) monitoring and control of learning strategy, and (3) reaction and reflection. Meanwhile, Winne and Hadwin (1998, 2012) divide Pintrich's first phase into task definition and goal setting and planning. Therefore, self-regulated learning occurs in four basic phases, (1) task definition, (2) goal setting and planning, (3) studying tactics, and (4) adaptations to metacognition. This model hypothesises that information processing occurs in every phase of the model as an interaction of students' conditions, operations, products, evaluations, and standards (COPES). This indicates that there is a product in every phase of learning. For example, the product of task definition was the characteristics of the task while the products

of goal setting and planning are learning goals and the plans to achieve the goals (Winne & Hadwin, 2008).

In the context of computer-based learning environment (CBLE), Azevedo, Johnson, Chauncey, and Burkett (2010) suggested that regulating one's learning in CBLE can involve 1) analysing the learning context, 2) setting and analysing meaningful learning sub-goals, 3) determining which learning and problem-solving strategies to use, 4) assessing whether selected learning strategies are effective in meeting the sub-goals, 5) monitoring and making an accurate judgement regarding one's emerging understanding of the topic and contextual factors, and 6) determining whether there are aspects of the learning context that can be used to facilitate learning.

In phase 1, learners are expected to understand the learning context, which can include the course detail and the learning environment. This process results in the understanding of the task expectation and the possible supports from the environment. Winne and Hadwin (1998, 2012) refer this phase as forethought phase where prior knowledge, of the content and the metacognitive knowledge, is one of the success determinants. In the CBLE context, prior knowledge of the learning environment will also be essential as it can help students focus on the accurate materials, which will lead to an appropriate sub-goal setting, and determine effective strategies to learn the contents. As an example, a student knowing that they lack vocabularies to do a speaking task will appropriately set the correct goals of their learning. The student will then decide which strategies to use to accomplish their goals. In contrast, not knowing their prior knowledge will result in the double curse of incompetence (Loibl & Rummel, 2015); the failure to set the correct sub-goals and to choose the efficacious learning strategies.

2.7.1. SRL and academic achievement

The majority studies on the relationship between SRL and non-academic outcomes reported the influence of SRL on students' satisfaction, engagement, interaction, and perceived learning outcome. The SRL aspects in those studies covered resources management (Dumford & Miller, 2018; Mikum, Suksakulchai, Chaisanit, & Murphy, 2018; Noh & Kim, 2019; Wichadee, 2018), motivational beliefs (Goda et al., 2015; Mikum et al., 2018; Paechter & Maier, 2010; Yamada et al., 2016), cognitive engagement (Cacciamani, Cesareni, Martini, Ferrini, & Fujita, 2012; Pellas, 2014; Yamada et al., 2016), and metacognitive knowledge (Cacciamani et al., 2012; Dumford & Miller, 2018; Pellas, 2014). The aspects of SRL on satisfaction, motivational beliefs (including self-efficacy), and resource management (including peer interaction, help-seeking, learner-instruction

interaction, time management) were reported to influence students' satisfaction on the learning program. Engagement to the learning material and interaction among students and between students to the teacher was influenced by self-efficacy and metacognition, goal orientation, motivation, and help-seeking. These studies also reported the contribution of SRL to the students' perceived learning outcome by considering their learning activities. In conclusion, having good resource management, motivational belief, and metacognitive knowledge help students to have better non-academic outcomes. The next question is whether it influences the academic outcome.

The studies on the influence of self-regulated learning on academic outcome have been examined with positive results. Two meta-analyses of studies on self-regulated learning showed high effects of SRL on academic success (Dignath & Büttner, 2008; Ergen & Kanadli, 2017). The moderating factors in these studies included learning strategy, course type, school level, research design, resource management, and metacognitive strategies (Ergen & Kanadli, 2017). In terms of the intervention, the teacher-researcher who gave motivational and reflective strategies was a more effective intervention for academic achievement compared to the classroom teacher (Dignath & Büttner, 2008). In conclusion, the majority of studies on SRL revealed the positive contribution of SRL on achievement with various moderating factors so that alternative strategies to enhancing SRL need investigating.

2.7.2. Interventions in SRL skill development

There have been ample interventions to improve students' self-regulated learning skills in CBLE, and an animated or computerised tutoring agent is one of the popular ones. Animated tutoring agents are animated characters that help students learning like a tutor—based on the definition of animated pedagogical agents in Craig, Gholson, & Driscoll (2002). One of the reasons to employ animated or computerised tutoring agents is the rare success of a human tutor to help students with deep conceptual understanding and the development of sophisticated self-regulatory skills (Chi, Roy, & Hausmann, 2008; Graesser, Mello, & Person, 2009; Graesser & McNamara, 2010). In addition, it is a challenge to train human tutors who can be consistent, precise, complex, adaptive, and durable to build effective interaction-centred tutoring (Graesser & McNamara, 2010). Thus, animated agents have been chosen by researchers as there is no significant difference between human and animated tutor (Graesser, Jeon, & Dufty, 2008; Graesser, Penumatsa, Ventura, Cai, & Hu, 2007; VanLehn et al., 2007). Due to the superiority of animated agents, studies on the use of these agents to help SRL development have been carried out.

MetaTutor is the most comprehensively studied computerised pedagogical agent to support self-regulation skills in the learning events (Azevedo & Johnson et al., 2010; Azevedo, Johnson, Chauncey, & Graesser, 2015; Azevedo et al., 2012; Graesser & McNamara, 2010; Trevors, Duffy, & Azevedo, 2014). In a long project of supporting students' self-regulation, Azevedo and colleagues designed MetaTutor (e.g., Azevedo, Moos, Johnson, & Chauncey, 2010; Greene & Azevedo, 2009; Taub, Azevedo, Bouchet, & Khosravifar, 2014; Taub et al., 2019), an integrated online tutoring system. A machined tutor was embedded in the system to help students with their self-regulation—in all phases of self-regulated learning (Azevedo & Moos et al., 2010). Multi-channelled data collecting techniques were used—for example, eye-tracking, think-aloud method, and log files—to capture cognitive and meta-cognitive strategies during the learning process. Results suggest that employing MetaTutor helps students with their self-regulated learning skills, more specifically on processes related to planning, metacognitive monitoring, learning strategies, and methods of handling task difficulties and demands.

This research predicted that the mechanisms of PF would affect the number of SRL strategies. In the MetaTutor, the moderating factors anticipated to affect students' SRL skill development included prior knowledge and adaptive scaffolding (Taub et al., 2014; Trevors et al., 2014), prompt and feedback (Azevedo et al., 2012; Duffy & Azevedo, 2015), and negative emotions on the accuracy of SRL behaviour (Taub et al., 2019). Prompts and feedback affected the number of observed SRL strategies, while prior knowledge and adaptive scaffolding influenced the quality of note taking. It is interesting that negative feedback could affect the quality of SRL strategies differently. Surprise negatively influenced the accuracy of metacognitive judgement, but frustration positively affected the accuracy of note takingnote taking (Taub et al., 2019). The next paragraphs covered the focus of the SRL strategies in this research.

This research focused on monitoring and learning strategies, given that these two aspects are essential in the development of self-regulated learners (Winne & Perry, 2005). Monitoring is the entry to SRL (Winne & Perry, 2005), as without which evaluation to one's learning would not happen. The products of the monitoring process led to the choice of effective learning strategies. Learning strategies are tools to approach the information in the instruction sessions or task completion. In addition, studies have reported the effect of PF on students' prior knowledge activation (Kapur, 2008, 2014b, 2015, 2016; Kapur & Bielaczyc, 2012) and knowledge gap recognition (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), which may direct the students' choice of learning strategies to use

in the instruction. During the consolidation in PF learning design, comparison between students' solution and canonical solutions (as the standard) will likely support monitoring students' learning performance. Therefore, both monitoring and learning strategies are worth researching in relation to the application of PF learning design.

This research looked at the process data of PF and DI students—video captured—when they watched the instructional video during the consolidation phase. Their responses toward the instructional video were coded based on Greene and Azevedo's (2009) coding schema. Emphasis was mainly put on the motivation, strategy use, and metacognitive monitoring—macro or micro levels. The PF mechanisms support the development of SRL strategies, which will be the focus of the next section.

2.7.3. Self-regulated learning and Productive Failure

Studies on PF in relation to self-regulation is still in its infancy, in which the results were students' perspectives not observational. In Kapur & Bielaczyc (2012) and Jacobson et al. (2015), for example, students claimed to engage and develop their metacognitive and self-regulatory skills while participating in a PF-based learning activity. In their research on designing for Productive Failure, Kapur and Bielaczyc found that a number of participants noted the opportunity to develop their self-regulation during the discussion. Similarly, students engaged deeper with the learning model (Pathak et al., 2011), and they asked more questions in the consolidation phase (Jacobson et al., 2017). The cognitive mechanisms of PF – prior knowledge activation and differentiation, knowledge gap recognition, and schema abstractions (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2015) – were predicted to affect PF students' dynamics during the consolidation phase. It might affect (1) the goal setting and planning, (2) study tactics, and (3) monitoring. The following paragraphs elaborate on the connection between SRL and PF.

First, the goal setting and planning in the first SRL phase might be facilitated by the prior knowledge activation and knowledge gap recognition. The initial problem-solving in PF enables prior knowledge activation and differentiation (Kapur, 2008, 2014b, 2015, 2016; Kapur & Bielaczyc, 2012) and general knowledge gaps (Loibl & Rummel, 2014a). Within the SRL theorising, prior knowledge may affect the accurate goal setting in SRL when combined with current learning context (Pintrich, 2000) and the quality of note taking (Trevors et al., 2014). Knowledge gap recognition, which was facilitated during the first problem solving and the consolidation (Loibl & Rummel, 2014a), is predicted to help set personal goals as well. Once prior knowledge is activated and differentiated during the idea generation and exploration, students' position relative to the target knowledge is located.

Therefore, it helps students set their personal goals in the planning stage SRL (Winne, 2017b; Winne & Hadwin, 2008). English as a Foreign Language Learners (EFL), for example, might understand their interlanguage position (vocabulary, grammar, and pragmatic knowledge and skills) relative to the standard of skills needed to finish the task after finishing the first task prior to instruction. Therefore, setting their personal goals, in which aspects they should focus on in the subsequent instruction, should be easier.

Second, PF helps students determine effective learning strategies during the instruction. It is evidenced that solving novel problems prior to instruction triggers the acknowledgement of general knowledge gaps (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015) and comparing and contrasting students' solution to the canonical one helps acknowledge specific knowledge gaps (Loibl & Rummel, 2014a, 2015). It is hypothesised that understanding specific gaps prior to instruction can help students determine which learning strategies to use to fill the gaps. Once prior knowledge is activated, and knowledge gaps are acknowledged from the comparison of the prior knowledge and the expected standard, setting personal goals (or sub-goals in SRL) will likely be easier. Therefore, there is a potential that PF students excel more useful learning strategies provided they have understood their knowledge gaps accurately.

Third, monitoring during the learning course is more likely anticipated in the PF given that students have known (or even experienced finishing the task based on) the expected standard during the initial problem-solving. Monitoring is carried out by comparing the current condition against the pre-existing standard (Winne, 2017b; Winne & Hadwin, 2008; Winne & Perry, 2005). The first monitoring opportunity might happen when PF students solve a new problem prior to instruction, which enables PF students to monitor their prior knowledge (Kapur, 2015; Kapur & Bielaczyc, 2012) relative to the standard (the problem). The next monitoring opportunity could occur when the explicit instruction compares and contrasts between the students' solution and the canonical solution, which is prescribed within the PF design (Kapur & Bielaczyc, 2012; Sinha & Kapur, 2019). The canonical solution can be used as the standard of the monitoring process. In DI, on the other hand, as the canonical solution has been presented before the problem solving, students might experience false mastery that they do not recognise their gaps. As a result, the monitoring strategy might not be used due to the false mastery feeling.

As the correlation is a prediction and little PF-related studies have been investigated on self-regulated learning, this research compared SRL strategies among PF and DI students. Provided that self-regulation is an essential determinant of student's success, this construct is

worth investigating in the PF-based design. This research focused more on monitoring and learning strategies provided they are essential characteristics of highly self-regulated learners (Pintrich, 2000; Winne & Perry, 2005). While there have been little investigations on the use of PF in language learning, this research is expected to fill the gap by employing PF in foreign language learning with a special observation on students' behaviours during the explicit instruction in the consolidation phase. As students in the PF group activate their prior knowledge, experience failure, and experience complex real-world problem longer, it is hypothesised that more self-regulated learning strategies are enacted during the consolidation phase by PF students.

2.8. Research Questions and Hypothesis

Foreign language learners, including Indonesians, find speaking in the target language a challenge. This is partly due to their lack of vocabulary and syntax knowledge (Cahyono, 2008; Kim, 2020). Studies on task-sequencing within Task-based Language Teaching (TBLT) suggested simple-complex task sequencing (e.g., Allaw & McDonough, 2019) – very little suggested complex-simple task sequencing (e.g., Baralt, 2014) – to successfully help learners improve their knowledge and skill in the target language. Drilling practices, however, still dominated the application of TBLT in the EFL context (Criado, 2013; Kumaravadivelu, 2006; Littlewood, 2013). In addition, the explicit instruction prior to task completion – regardless of the task sequence – restricted students' creativity due to too much reliance on the model to finish the task (Ellis, 2018).

Productive Failure, a learning design that allows students to solve a new problem before explicit instruction (Kapur, 2008; Kapur & Bielaczyc, 2012), was proposed as a potential instructional strategy to teach grammar ability. The proposed research focused on answering the following questions:

1. Is there a different achievement in the written description of places between the experimental (PF) and control (DI) groups?
    a. Overall written description
    b. Declarative knowledge
    c. Procedural knowledge
2. Is there a different achievement in the oral description of places between the experimental (PF) and control (DI) groups?
    a. Overall Oral description
    b. Declarative knowledge
    c. Procedural knowledge

3. Are there different behaviours during the instructional video between the experimental (PF) and control (DI) groups from self-regulated learning theorising?

For the first research question, the project limits the sub-skills of vocabulary and syntax in written production as studies suggest that they are the sub-skills that are measurable within a short period of treatment (Saito, 2014; Yoshii & Flaitz, 2002) and have been a challenge for foreign language learners, Indonesian students (Cahyono, 2008; Kim, 2020). It is hypothesised that PF students will do equally to DI students in declarative knowledge (vocabulary). Research to date found that there is no difference in the development of declarative knowledge (vocabulary in this research) between PF students and DI students in their science classes (Jacobson et al., 2017; Lai, 2017). As students from both groups would have relevant prior knowledge about vocabularies (things at the beach and spatial preposition), they should be able to learn additional declarative knowledge equally. For example, students from both groups have known a preposition "on" and "above." This prior knowledge should be helpful for them to learn a preposition "on top of" by associating the unique location and compared to the previously learned location with a similar preposition. Both groups should equally succeed in the new preposition associated with a picture representing the position.

In the syntactic knowledge (procedural knowledge), we hypothesised that PF students would do better in the post-test targeting syntactic accuracy in the written description of places. Previous studies on PF found that PF students do better on procedural knowledge (Kapur, 2012, 2014a, 2015; Kapur & Bielaczyc, 2012). By comparing both solutions, students will understand the critical features of the targeted content, which will lead to better performance in procedural knowledge (DeCaro & Rittle-Johnson, 2012). Once they understand the critical concept, they can use it in the written description of places after the instruction.

In addressing the second research question, the project still limits the sub-skills of vocabulary and grammar used in the oral description of places with the same reason (Saito, 2014; Yoshii & Flaitz, 2002). It is hypothesised that students from both groups will do equally both in the lexical and syntactic accuracy of the oral description of places. The reason is that oral production is a complex system (Bock & Levelt, 1994; Levelt, 1999, 2001) and students' oral performance is dependent on the frequency of practice (Bygate, 2013; Ho, 2017; Kim & Tracy-Ventura, 2013). Both PF and DI groups got the same number of practices despite the arrangement. However, if we take Loibl and Rummel's (2014a, 2014b) notion that procedural fluency is depending on the number of practices after the instruction, DI group

would have better performance on lexical and syntactic accuracy in the oral description of places. This is in line with studies in language learning that post-instruction practices play an important role in determining the success of the use of newly-learned expressions in the oral production of the target language (e.g., Finardi, 2008; Fukuta, 2016; Ho, 2017; Jacoby, 1978).

For the third research question, the hypothesis is that PF students would have higher self-regulated learning skills due to the cognitive mechanisms of PF pedagogical strategy. As PF students had the opportunity to activate their prior knowledge and acknowledge knowledge gaps (Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a), they would likely set their personal sub-goals more accurately and plan their learning execution better. Accurate personal sub-goals and knowledge gaps recognition allowed PF students to own more prerequisites to choose which learning strategies to employ during explicit instruction. Effective learning strategies would likely succeed to fill their learning gaps. When accurate personal sub-goals and knowledge gaps are at hands combined with the canonical solution as the learning standard, monitoring learning should happen smoother. Therefore, it is hypothesised that PF students may enact more learning strategies and metacognitive monitoring.

Chapter 3. Method

This project compared students' lexical and syntactic accuracy in students' written and spoken descriptions of places between students in *Productive Failure* (PF) and *Direct Instruction* (DI) groups. Both groups carried out the same activities but in a different pedagogical sequence of the tasks and instruction. After the first task, PF students watched the instructional video, while DI students watched the instructional video before the first task. This chapter describes the research methods used in this project.

This research aims to compare two pedagogical strategies in teaching English as a foreign language, especially on the topic of describing places. This study specifically answers the following questions:

1. Is there a different achievement in the written description of places between the experimental (PF) and control (DI) groups?
    a. Overall written description
    b. Declarative knowledge
    c. Procedural knowledge
2. Is there a different achievement in the oral description of places between the experimental (PF) and control (DI) groups?
    a. Overall Oral description
    b. Declarative knowledge
    c. Procedural knowledge
3. Are there different behaviours during the instructional video between the experimental (PF) and control (DI) groups from self-regulated learning theorising?

3.1. Research Design

This study employed an experimental design (Creswell, 2012; Creswell & Creswel, 2018) with a hybrid data collection procedure (Phakiti, 2014) to investigate these research questions. It was experimental in that the treatment was given in a lab-like environment, and samples were drawn randomly. The advantages of the experimental design are, among others, a high level of control, specific conclusions, and a possibility of duplication (Gaille, 2017). The data collection was hybrid as both quantitative and qualitative data were gathered (Phakiti, 2014). This research's quantitative data were students' vocabulary and syntax of written and spoken description of places from the pre-and post-tests. In contrast, the qualitative data were students' behaviours observed during their attendance while watching

the instructional video. Figure 2 describes the process of the research, which is described in the next sections.

3.1.1. Participants

One hundred and twelve students (83 females and 29 males) from three universities in Yogyakarta Special Region (*Daerah Istimewa Yogyakarta* - DIY) participated and were included in the analysis. One hundred and fifty-five students from three universities in Yogyakarta Indonesia expressed their interest to participate in this research after the researcher presented the research in their classes. They joined the *SL* training, and only 133 participants consented on the second day of the training. They were randomly assigned to the experimental and control groups using www.randomizer.org. Only 113 were compliant to the instructions (56 DI and 57 PF groups), who did all the activities assigned to them. The non-compliant participants (N=20) did not do at least one activity assigned to them. Non-compliant students were not included in the analysis to ensure the accuracy of the results. One participant from the PF group was randomly removed from the study to maintain an equal number of participants in both groups. Excluding one random participant out of 57, regardless of the response, will not affect the results as it is only slightly more than 1%.

The purposive voluntary sampling technique (Jupp, 2006) was used. It was purposive in terms of universities' choice and voluntary in terms of students' participation from each university. Three universities in Yogyakarta Special Region were chosen to assume that they represent Indonesian English language education (ELE) department students given that their students come from all provinces in Indonesia. ELE students from the three universities participated voluntarily in this research after the information sessions in their classes, in which the faculty permission has been granted. Thus, there was no pressure for the students of the three universities to participate.

The sample size was determined based on the previous studies in PF, with the smallest sample size of 75 in quantitative research by Kapur (2014b). One hundred and twelve ELE students took part in this research. They were 28 from Yogyakarta, 47 from other areas in Java, and 37 from other islands, a good representation from all parts of Indonesia. Despite the geographical representation, this study was not expected to be generalised to Indonesian ELE department students due to the small number of samples compared to the ELE department freshmen population.

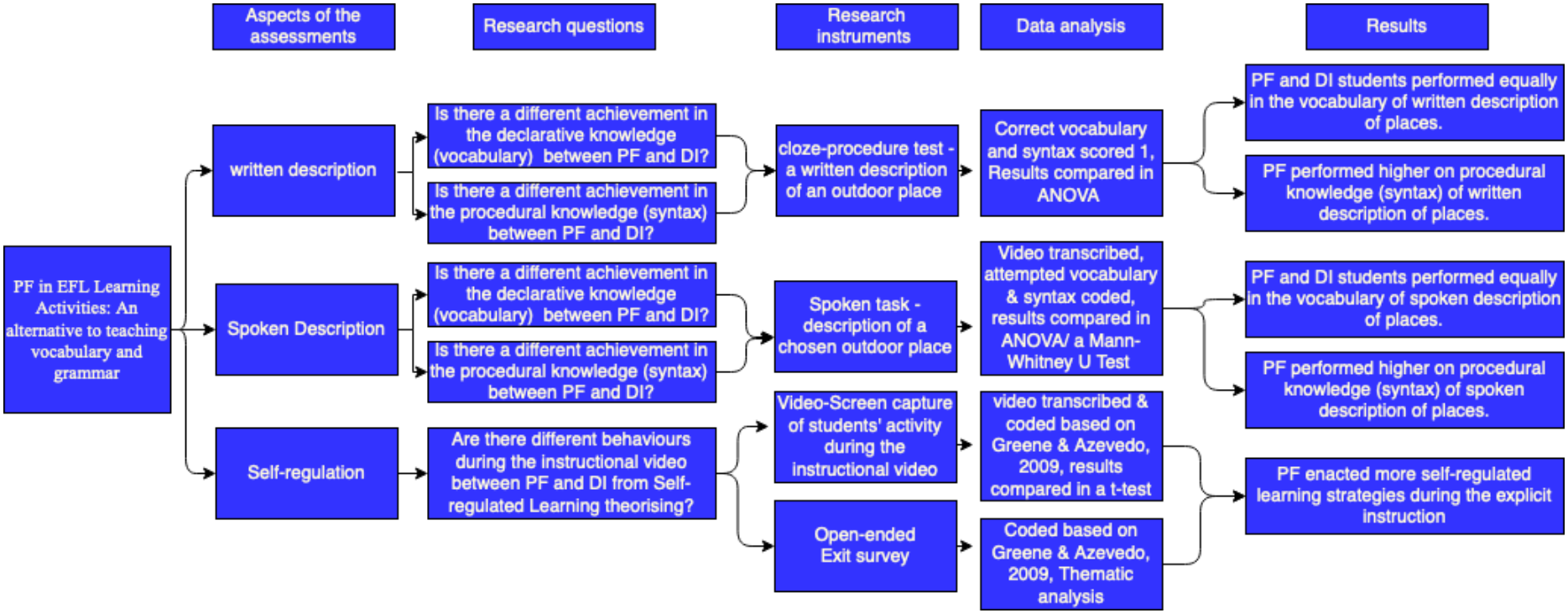


Figure 2 The flowchart of the research process

They have been studying English in the classroom basis for at least eight years prior to their participation in the project. They belong to EFL learner groups, who can understand sentences and frequently used expressions related to areas of most immediate relevance, for example, basic personal and family information, shopping, local geography, and employment. They can communicate in simple and routine tasks requiring a simple and direct exchange of familiar and routine information. They can describe, in simple terms, aspects of his background, immediate environment and areas of an immediate need as well (Council of Europe, 2001). Table 2 depicts the participants' demographical characteristics based on region, institution, pre-test score, and the statistical test results.

Table 2 Demographical Characteristics of Participants

		PF (%)	DI (%)	Total (%)
Origin	Yogyakarta	18.75	22.32	41.07
	Other cities in Java	12.50	11.61	24.11
	Out of Java	18.75	18.00	34.82
SL experience	Yes	8.93	12.50	21.43
	No	41.07	37.50	78.57
Live-abroad experience	Yes	0.89	0.89	1.79
	No	47.32	48.21	95.54
English Learning Experience	Formal school and Learn at home	6.25	7.14	13.39
	Formal School and English Courses	25.00	23.21	48.22
	Formal school	18.75	19.64	38.39

3.1.2. Research settings

The language laboratory of English Education Department, Universitas Islam Indonesia (Figure 3), was the setting of the project. Students carried out the activities from fifteen computers connected to high-speed internet from Eduroam and UIIConnect Wifi channels. A screen capture software named *Bandicam* and a webcam were used to capture the students' activities and facial expressions.

This research's virtual setting was *SL* Indonesia headquarters, established by a group of Indonesians playing *SL*, intended to introduce Bahasa Indonesia to *SL* users, who are non-Indonesian language speakers. Two areas of the *SL* Indonesia Headquarter were used. The bottom level (figure 4a) was used as the landing space where participants arrived and chatted with the conversation companions and other *SL* users. The upper level (figure 4b) was set as the main venue where participants could access the instructions and played the instructional video. In addition, Virtlantis beach (figure 4c) and NCI beach (figure 4d) are *SL* sites used

when participants were doing their tasks. They were the beaches to describe, one of which was set as a virtual conference venue while the other one was a typical beach.

Figure 3 The site of the data collection (Language Laboratory)

Figure 4 Virtual places used for a) the first communicative task and model description, b) the main venue, c) the welcoming area, and d) the second communicative task

Virtual worlds are "shared, simulated spaces inhabited and shaped by their inhabitants who are represented as avatars. These avatars mediate our experience of this space as we move, interact with objects and interact with others, with whom we construct a shared understanding of the world at that time" (Girvan, 2018, p 1099). *SL* is one of the virtual worlds established by Linden Lab. The virtual world *SL* was chosen provided it was free and available on multiple platforms. It was one of the most popular virtual worlds available

(Dickey, 2011) offering a rich diversity of contents within an immersive environment. The visual nature and interface added to the immersive feeling and created more realistic situations for communication. The number of users and locations (islands) provides the students with plenty of rooms to experiment and opportunities to interact with native speakers. *SL* also integrates both voice and text chat technologies so students can speak in real-time voice. This technology is not available in many other virtual worlds.

3.1.3. Activity sequence of PF and DI group

While this research utilised MUVE, the focus is not on the technology but the pedagogical strategy. This research emphasised how students effectively learn in technology-based language learning as the same MUVEs were used in both groups. The difference was on the pedagogical strategy, PF and DI. Unlike other studies on MUVEs, this project was designed for individual learning using technology. It investigated how students could benefit from the design, especially how it affected their monitoring and strategy tactics in independent learning.

This research was designed based on PF (Kapur, 2008; Kapur & Bielaczyc, 2012) as the experimental condition and DI (Klahr & Nigam, 2004) as the control condition. Both groups carried out the same activities with the difference in the activities' sequence (see Figure 5). The PF allows problem-solving prior to instruction while DI lets students do the problem solving after the instruction. The research's PF design scored 4 out of 5 in the fidelity check (for a complete description, see Sinha & Kapur, 2019). The design in this research failed to comply with the collaborative nature of the problem-solving in the PF design. Our purpose was to facilitate personalised online learning, so failing in the collaboration nature would still be acceptable.

Relevant to PF design, the design comparison of the research is depicted in Figure 5. The emphasis for the treatment was in the blue box – communicative task and explicit instruction. There were two phases in the PF design; (1) idea generation and exploration and (2) consolidation and knowledge assembly (Kapur, 2008; Kapur & Bielaczyc, 2012). In the first phase, students described a place to a conversation companion based on a communicative task in Figure 7. They have not learned about the task before. PF students watched an instructional video in the consolidation and knowledge assembly and did the second communicative task in Figure 8. The DI students, in contrast, watched the instructional video before doing the first communicative task (Figure 7). After that, DI students did the second communicative task (Figure 8) just like the PF students. The second task was describing a similar place in *SL* named the NCI beach to the conversation

companion. The second task was designed to facilitate the knowledge assembly and the development of procedural knowledge. In line with Loibl and Rummel's (2014a, 2014b) suggestion, more practices after the instruction should be provided, especially for PF students, if procedural knowledge is to be tested. Before the main comparison, both PF and DI students joined in training on *SL* and took a pre-test comprising a written (Appendix 1) and spoken description of places (Appendix 2). A complete description of the activities is discussed in the following sections.

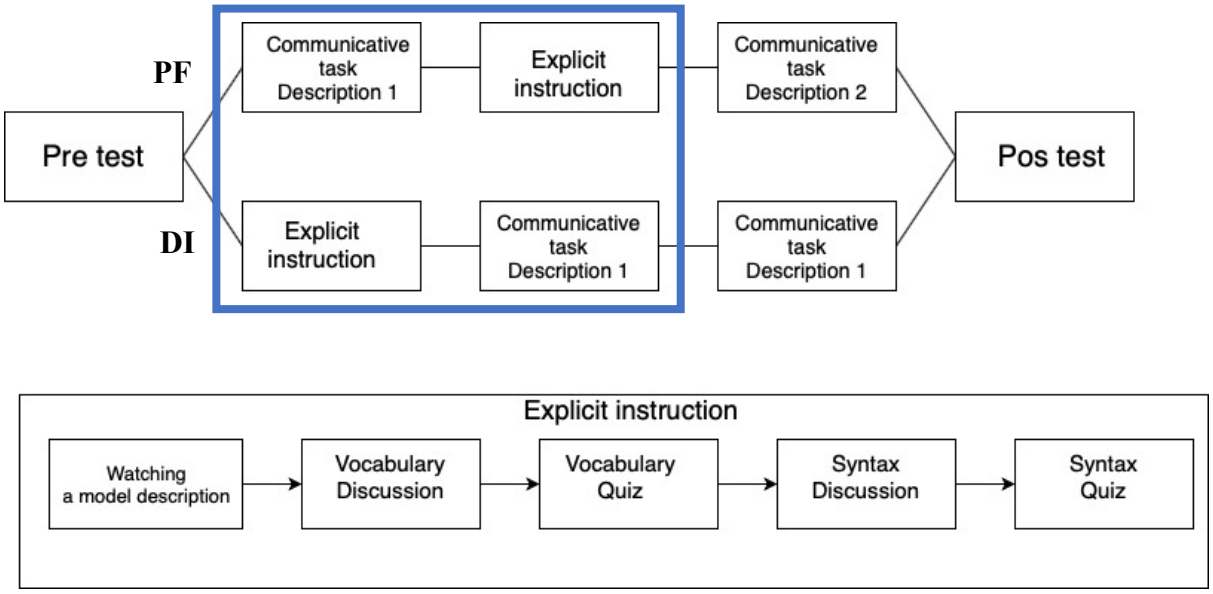


Figure 5 The design comparison of PF and DI

3.2. Data sources

Multiple data were collected before and after the treatment, both quantitatively and qualitatively. This research is a hybrid experimental (Phakiti, 2014) study in which quantitative measurement was the primary data, supported by the qualitative ones. The data sources included students' learning outcomes, students' experience in learning English, and a recording of students' behaviour during the video viewing.

3.2.1. Students' learning outcomes

The students' learning outcome was assessed using pre-and post-tests (Appendix 1&2) targeting participants' lexical and syntactic accuracy in written and spoken description of places assigned to the students from both groups. The lexical accuracy measured the declarative knowledge while the syntactic accuracy tested the procedural knowledge, the ability to correctly apply the learned procedure (Rittle-Johnson & Schneider, 2015), in written and oral description of places.

The pre-and post-tests were the same tests with two sections, a cloze-procedure test and an oral description task. The written test asked the participants to complete sentences used in describing places targeting the vocabulary and grammar used in describing places. There were five vocabulary items and 11 grammar items. The performance test, on the other hand, asked the participants to choose one out of three virtual gathering venues and describe it to the conversation companion. The students' performance in both pre-and post-tests were recorded in video format using *Bandicam*. They named the file with their *SL* name and saved the data into the local computer storage.

The written assessment was a researcher-made test based on Bachman and Palmer's (2010) frameworks for constructing a language test. It targeted the vocabulary and syntax used in describing places. It was initially a cloze-procedure test with 16 blank parts from a paragraph describing a virtual beach called Virtlantis. As students did the tests on a Google form, one test item was presented separately based on the sentences with the blanks with the surrounding sentences. It was scored binomially in which correct answers were rewarded "1", while incorrect answers were scored "0." The reliability of the items was acceptable with the KR-20 coefficient of .788 (Salkind, 2010).

Figure 6 Examples of places to describe in the pre-and post-tests

A spoken assessment (Appendix 2) measured the participants' lexical and syntactic accuracy in describing outdoor places in *SL* orally. Students, represented by their avatars, described one out of three sites in *SL*. This is a performance-based test carried out in *SL*

before and after the treatment as the pre-and post-tests. They teleported to one of the places in Figure 6 and described it to the conversation companion representing remote audiences. Screen capture in video format was taken from the students' computer. These videos were then verbatim transcribed for the data analysis.

3.2.2. Students' background

An open-ended pre-survey (Appendix 3) was used to assess the participants' demographic background and their online and English learning experiences. The survey found that participants were 18 – 20 years old and from nearly all areas in Indonesia. They have learned English both formally at school and English courses or informally via interaction with relatives. The complete demographic profile of the participants is in Table 2.

3.2.3. Students' perceptions

An exit survey (Appendix 4) was assigned to the participants via google form right after they finished the last activity. Nine open-ended questions were presented to the participants about their perceptions of their learning experience via *SL* and the assigned pedagogical strategies. The survey was open for the participants to answer the questions, and its results were used to double-check the data from the video observation. The students' answers from the exit survey (for sample, see Appendix 12) was used to confirm the result of the observation (for sample observation sheet, see Appendix 14).

3.2.4. Process data

To answer research question 3 on the students' behaviour (their learning strategies and monitoring process) during the consolidation phase, a screen recording was carried out as a substitute of classroom observation to capture students' behaviour. The behavioural measure (systematic observation) was chosen instead of a self-report questionnaire to ensure the accuracy of students' SRL behaviours as suggested in a recent meta-analysis on SRL measures (Rovers et al., 2019). The advantage of measuring a specific cognitive process during the actual learning (Greene & Azevedo, 2007) was the other reason for using video capture to gather the data. Participants worked on an internet-connected desktop with an installed screen recorder, *Bandicam*, to record students' activities during their participation in the research. A camera was also attached to the computer to record the students' facial expression. Recording students' online activities allowed high control of the treatment implementation, which enabled the researcher to check the participants' compliance toward the instruction by seeing the video recordings.

The use of *Bandicam* and camera allowed the researcher to capture students' behaviour during the consolidation phase. A capture on their focus from the instructional

video and what created excitements from the description model was captured using the software. Errors in the first task and excitements in the video model could be interpreted as both gaps and individual learning goals. During the consolidation, the video records were then expected to reveal data on their focus and their responses to the instructional video interpreted as cognitive learning and metacognitive monitoring strategies based on Greene and Azevedo's (2009) coding protocol on self-regulated learning.

3.3. Data Preparation and Analysis

The data were prepared carefully to ensure accuracy of the data entry and the analysis; a PhD student rechecked the data entry. The next section describes the data preparation and analysis method of the research.

3.3.1. Data preparation

To guarantee the scoring's objectivity, the answers of the written assessment (Appendix 5) were filled into an online sheet and then scored blindly before being separated into PF and DI. The first grouping was based on the assignment prior to the data collection. Meticulous checking was carried out by watching all videos to ensure that the participants carried out all the activities in the correct sequence. Correction on grouping was done as a number of students did not do the activities as assigned. Initially assigned PF students were moved to the DI group provided they did the activities in the DI sequence, and vice versa.

To ensure the correct data entry, a PhD student double-checked the accuracy of the data entry. The error rate in the data entry was 2.25% in those three out of 133 participants were entered incorrectly. The spoken pre-test or post-test from the three participants were entered into another participants' slot. A correction was done by moving the data to the correct participant.

The students' oral pre-and post-test were verbatim transcribed (Appendix 6) by two last semester English students. All words in the video were transcribed regardless of the repetition (Paulus, Lester, & Dempster, 2014), was carried out by two English language education graduates, professionally recruited for the transcribing. They worked on the same files for nearly 20% of the participants with the similarity index of 89% counted in word unit. They transcribed the rest of the data independently. The first transcriber worked on the 53 files, while the second transcriber worked on the records from 54 participants.

3.3.2. Data analysis

To answer research question 1, a written test with items targeting vocabulary and syntax in describing places was used to capture the participants' declarative (vocabulary) and procedural (syntax) knowledge of a written description of places. A between-subject effect in

a repeated measure ANOVA was used to determine the impact of the treatment on the written vocabulary and syntax between the control and experimental groups.

For research question 2, videos of the participants' verbal description were verbatim-transcribed (Paulus et al., 2014). Attempted vocabularies and syntax were recorded based on their occurrence instead of the frequency. One word attempted five times in a spoken description, for example, was scored "1" not "5." A between-group effect within a repeated measure ANOVA was used to compare the impact of the treatment on the students' oral vocabulary and syntax in describing places between PF and TBLT/DI groups.

For research question 3, students' videos on their activities during the consolidation were analysed based on coding for self-regulated learning behaviours (Greene & Azevedo, 2009). The coding protocol (Table 3) defined students' behaviours in accessing the instructional video, i.e., whether they finished the video, how they responded to the teacher's questions and instructions, the reaction to their gap-filling activity, and how enthusiastic they were toward the video. The categories were then matched with Azevedo's self-regulation cycles (Azevedo et al., 2010; Greene & Azevedo, 2009). Statistical comparison between groups was carried out by using t-test.

3.4. Learning Materials

The learning material in the consolidation phase was developed based on fluent English speakers' description models. Five English native speakers living in Australia described the same pictures, in which their common expressions were used in the description model to be used in the consolidation phase. This model was used as the standard to evaluate the students' description in the first challenge. However, description standards for the second challenge and the pre-and post-test used Indonesian fluent speakers' description. Common expressions of that description were used to assess the students' description. The tasks were considered complex based on the students' language skills and Robinson and Gilabert's (2007) Triadic Componential Framework for Task Classification.

3.4.1. Challenge – monologic task (Task) 1

The first task (Figure 7) allowed the participants to describe Virtlantis (Figure 4c)— a venue for a virtual gathering—as a one-minute YouTube video. Initially, students should have gone to Virtlantis and described the place, but a video capture embedded in YouTube had been used provided the owner in SL had demolished it. The participants described the place to the conversation companion on a private call while watching the YouTube video. The conversation companions were not allowed to watch the video. They were expected to use the target vocabularies and appropriate syntax before the instructional video in PF or after the

instructional video in DI in the consolidation phase. PF students were expected to activate their prior knowledge with this activity as they had not learned about a number of target vocabularies and syntax of describing places. On the other hand, DI students practised what they have learned in the instructional video in this activity.

Table 3 Examples of the coding of the observation notes

Micro-levels	Description	Examples from the observation notes
Macro-level: Motivation		
Motivation +	Learners showed positive behaviour during the video viewing.	"eyes on screen He took notes He watched the video attentively, bright eyes, smiley face".
Motivation -	Learners showed negative behaviour during the video viewing.	"She played the video jumping from one scene to the other quickly. She watched the video while playing a pen. She watched the video while being sleepy. "
Macro-level: Learning strategies		
Memorisation	Learners try to memorise vocabularies and other information from the video	She repeated "stilt house" several times.
Imitation (draw)	Learners try to imitate vocabularies and expressions from the video.	"She ... while mimicking the vocabularies mentioned. She kept saying the words."
Interaction	Learners respond to the teacher's questions or statements.	"She answered the questions from the virtual teacher. She described the place as asked by the virtual teacher. "
Reviewing	Students review what they have learned from the video	When the video finished, she reviewed by murmuring.
Taking notes (Copying texts from the recording)	Copying the text from the instructional video.	... paused after "melodic ocean waves" and took notes. Notes taken on aspects of describing places.
Revisiting a section (re-reading)	Learners revisit sections of the video.	"She played back the video. ... repeated the video for difficult words..."
Macro-level: Monitoring		
Feeling of Knowing	Students are aware of having some understanding of the content of the video.	... nodding whenever she knew the introduced vocabularies...
Judgement of Learning	Students are aware that they do not know the things they watch.	He often yielded confirming that he has used wrong vocabularies in the first description.
Self-questioning		
Macro-level: Task difficulty and demands		
Help-seeking	Learner seeks help regarding the adequateness of his/her answer	When prompted to do the quiz, she asked the conversation companion.
Control of content	Using features of the video to enhance the viewing of the information	... She turned on the English sub-titles in minute 2.

3.4.2. Challenge – monologic task (Task) 2

In the second challenge (Figure 8), the participants were asked to describe NCI beach (Figure 4d)—a training centre for a building class in *SL*. Permission has been granted by the NCI beach owner, which was to be described as a virtually-outdoor seminar venue. Participants had to give suggestions if it was a good place for a virtual seminar as well. They described NCI beach to the conversation companion on a private. The participants were at the venue while the conversation companion stayed in *SL* Indonesia headquarter call to keep the description authentic. This activity was recorded in a video format.

We are going to have a gathering in *SL.* We would like you to recommend a place for a group of tourists. They are represented by your conversation companion. Please consider the followings when choosing the site:

1. Around 20 people are going to attend the gathering
2. The audience is already very busy in the real world. They need a relaxing place where they can listen to music or nature. They might want to dance over beautiful music.
3. If you are one of the participants, what do you like and what do not you like about the place?
4. Choose the starting point like the pictures. Say where you are and start describing.

Figure 7 Communicative task in the idea exploration and generation

3.4.3. Consolidation video

A pre-recorded instructional video was presented to the participants by accessing the link in a red ball in *SL* Indonesia headquarter (Figure 4b). The conversation companion can help them understand the instruction when needed. The consolidation video was 8:21-minute long, which is an ideal length, as suggested in Guo, Kim, and Rubin (2014) and Lin et al. (2017). In this video, an Indonesian female teacher presented the materials on describing places. Students could watch the video from the virtual display in *SL* headquarter, but nearly 80% of students watched it on YouTube. The instructional video was in three parts.

Please advise us!
We need a venue for an online conference on *SL.* We have a recommendation at NCI beach, but we are experiencing a technical difficulty. As we need to decide it in no time, we would like you to go and see the place. Please describe the place to us (represented by your conversation companion) if NCI beach is the right place for the conference. Please consider the followings when choosing the place:

1. Around 20 people are going to attend the gathering
2. The audience need a relaxing place while listening to speeches. The music of nature will be a good idea.
3. To make the conference look fine, participants need to sit on chairs while the speaker can stand in front of them. If you are one of the participants, what do you like and what do not you like about the place?

Choose the starting point like the pictures. Say where you are and start describing to your conversation companion.

Figure 8 Communicative task after the consolidation and knowledge assembly

The first part is a model from an L1 speaker's description was introduced to the students before the virtual teacher discussed how to describe a beach. The model description covered the vocabularies (things at the beach, prepositions, and related adjectives) and the

acceptable syntax used to describe a place. The target syntaxes were noun phrases and sentence patterns used for describing places. Illustrations were provided while the model description was playing by zooming or focusing the objects.

After the model description, a virtual female teacher discussed vocabularies and syntax used in describing places. Vocabularies were introduced from its form and meaning, followed by the example in sentence context in Production Based Instruction (PBI) (Shintani, 2015; Shintani et al., 2013). A verbal-visual word association strategy was applied (Myers & Chang, 2009), in which a related visual in a video was highlighted while a word was introduced. PBI was used as students were expected to use the vocabulary in both written and spoken production. Only essential vocabularies were emphasised provided the short period of the treatment. A quiz on Google form was presented for the students as a practice.

Along with the vocabulary, the syntax was also introduced deductively in PBI (Jean & Simard, 2013; Shintani, 2015; Shintani et al., 2013; Tammenga-Helmantel et al., 2014). After the syntax was used in the model description, formulas were presented along with the examples. A number of pauses, questions, and instructions were given after every material for practice or internalisation. Students were asked to do manipulation or creation tasks provided production-based instruction gives more durable productive knowledge (Shintani et al., 2013). This strategy was applied in this study with the manipulation task in the written assessment and creation task in the oral assessment to measure students' improvement in lexical and syntactic accuracy in the target language production. Table 4 gives an example of the grammar introduction. A quiz on Google form was given for the students as a practice. Appendix 7 describes the lesson plan for the consolidation.

Table 4 Types of Tasks in Production-based Instruction

Aspects of the production	Expected language expressions	Types of tasks	Examples
Written	On the left side of	Sentence completion	There are multiple tall, dark green palm trees ..... (1) of the image. (A picture of the place was attached)
Spoken	On my left side	Sentence production as a part of a complete spoken description.	Students were asked to describe a place from their first position after teleport. They might produce a sentence like "There are palm trees **on my left side**."

3.5. Conversation Companion

Seventeen conversation companions were recruited from LPDP mailing list, which members were postgraduate students from universities in English speaking countries.

Training on the use of *SL,* including simple troubleshooting procedures, were provided in three meetings on *SL.* The training was carried out until the conversation companions could handle all possible problems. When they continuously could not solve the technical issues, they were suggested to withdraw from this research voluntarily.

The volunteer's duties, among others, were greeting the participants, delivering instruction, and conducting an online interview. Volunteers greeted and chatted with the participants right after they logged in the *SL* for about 5 minutes. They were allowed to have an introduction and small talk about hobbies, but not about the research. Volunteers then invited the participants to the main venue and *SL* Indonesia headquarters to talk about the instruction represented in coloured balls. They discussed the instruction and let the students ask questions. Volunteers also delivered in-treatment instructions by a private call. Information for conversation companion is in Appendix 8.

3.6. Pilot Study

Fifteen participants from the same cohorts joined the pilot study held at the English Language Education Department computer laboratory, the same venue as the data collection. The pilot suggested a number of revisions were suggested from the first treatment design. Consequently, the following improvements were put in place.

1) Rules and transcripts for the conversation companion were provided. At about half of the conversation, the companions were tempted to scaffold the participants when describing places in the first and second task. It might affect the quality of the data due to the provision of help. A stricter rule was used for the volunteers with a consequence of not being used in the analysis if the helping practice remained.
2) The flexibility of the media from which participants should access the instructional video. Eight out of 15 participants found difficulties accessing the instructional video from the YouTube player inworld (in the Second Life world), which affected their emotion in learning. They suggested YouTube as an alternative to play the instructional video. Thus, participants could play the video either inworld or on YouTube.
3) The pilot suggested that the time allocated in each section was sufficient, especially on the first and second tasks. The participants had 2-3 minutes to describe the places.
4) Rewording the exit survey was needed to ensure the participants understand the meaning.

Chapter 4: Results

This research compares students' vocabulary and syntax achievements in describing places between Productive Failure (PF) and Direct Instruction (DI) groups and their behaviours during explicit instruction. More specifically, the research answers the questions on the different achievement between PF and DI students on 1) the vocabulary and syntax used in the written description of places, 2) the vocabulary and syntax used in the oral description of places, and 3) the different behaviour patterns during the instructional video between PF and DI groups.

This chapter describes the results of the study, presented under each research question. The summary of the descriptive results is presented at the beginning of every section. The result of the statistical analysis is then presented to answer the research questions. At the end of every section, a conclusion is drawn to prove the subsequent hypothesis. The first section describes the result of the written description of places.

4.1. Research Question 1: the Different Achievement in the Written Description of Places

The first research question of the study dealt with the differences in the learning gain of lexical and syntactic accuracy in the written description of places. Both groups underwent the same learning activities targeting description of places, but different in the sequence of activities. DI group watched an instructional video before the two tasks of describing places. In contrast, the PF group finished the first task before watching the instructional video, followed by the second task. Table 5 shows a summary of the written lexical and syntactic accuracy in the written description of places and the total of the written scores.

Table 5 Summary of the mean and standard deviation of the written test scores

Time	Lexical		Syntactic		Overall Written Test	
	Experiment	Comparison	Experiment	Comparison	Experiment	Comparison
Pre-test	3.38(1.88)	3.14 (2.19)	3.80 (1.58)	2.91 (2.47)	4.04 (2.35)	3.77(2.51)
Post-test	5.77(2.55)	5. 00(2.64)	6.39 (2.35)*	4.82 (2.68)*	7.07(3.18)**	5.66 (3.17)**

Notes: * significant at $p<.05$
** significant at $p<.001$

4.1.1. Learning gain on the overall written test.

To explore if both groups started from the same level of prior knowledge on the written description of places, an independent sample t-test was conducted before comparing the learning gain. The independent t-test revealed no difference between PF and DI groups on the written pre-test $t(110) = -.583$, $p = .56$, two-tailed. The two-tailed value was chosen as the direction could not be predicted.

Moreover, a one-way between-subject analysis of variance revealed no significant effect of gender F(1, 110), p = .088, two-tailed and origin F(2, 109) = .068, p = .934, two-tailed on the overall pre-test scores. There was, however, a significant effect of the students' pre-test based on institution F(2, 109) = 5.853, p = .004, two-tailed. Thus, the institution was put as a covariate in the mean comparison between PF and DI groups.

A mixed between-within subjects analysis of variance reported a significant main effect of pedagogical strategies on the students' learning gain of the written description of places F(1, 110) = 3.185, p = .039, one-tailed, η_p^2 = .028. A significant interaction between time and group F(1, 110) = 5.058, p = .013, one-tailed, η_p^2 = .044 was also reported. The result remained significant when the institution was controlled as a covariate F(1,109)=5.543, p=.01, one-tailed, η_p^2 = .044 with non-significant effect of institution F(1,109) = .574, p=.275, one-tailed, η_p^2 = .005. The statistical result means that the PF group outperformed DI group on the post-test of the written description of places given that the PF group's mean is higher than the DI group's (see table 5).

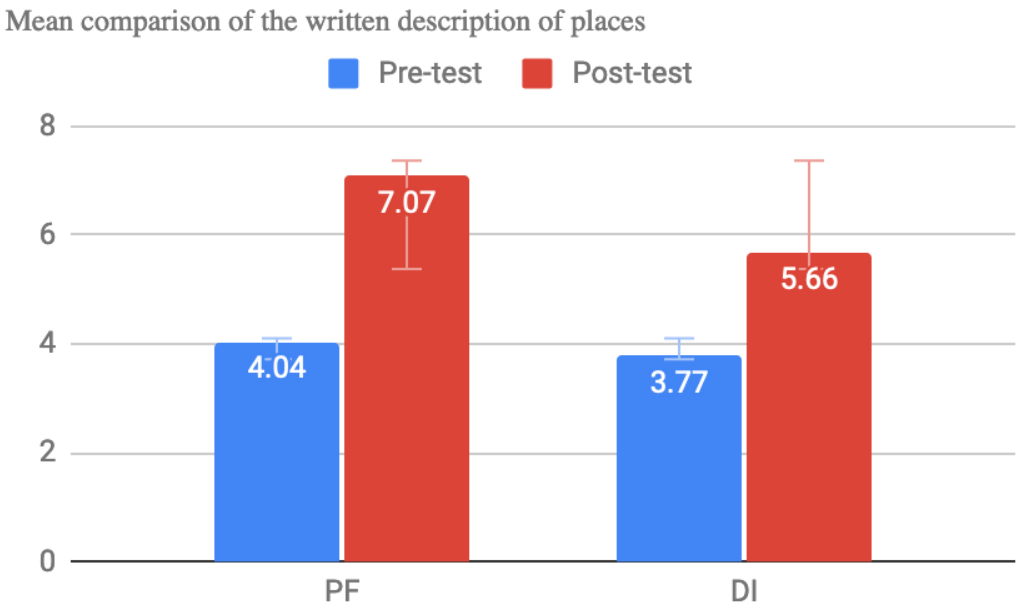


Figure 9 Learning gain on the written test

The result indicated that the hypothesis – there was no significant difference between PF and DI groups on the learning gain of the written description of places – was rejected in the one-tailed test. We concluded that students in the PF group performed significantly higher on the written description of places than those in DI group (see Figure 9).

4.1.2. Learning gain in the procedural knowledge of the written description of places.

PF and DI groups did not start from the same prior knowledge before the treatment provided a significant difference was reported on the pre-test of syntactic accuracy (procedural knowledge) in the written description of places t(110) = -2.277, p =.025, two-tailed. However, there was no significant effect of gender F(110) = 2.170, p = .114, two-tailed; origin F(2, 109) = .085, p = .918, two-tailed; and institution F(2, 109) = 1.265, p = .286, two-tailed on the pre-test scores on the syntactic accuracy in the written description of places. Therefore, a mixed between-within subject analysis of variance was employed to compare PF and DI's learning gain.

A significant between-subject effect of pedagogical strategies on the students' learning gain on the syntactic accuracy of the written description of places was revealed F(1, 110) = 10.640, p < .001, one-tailed, η_p^2 = .088. There was, however, no significant interaction between time and pedagogical strategies in within-subject contrasts F(1, 110) = 2.428, p = .061, one-tailed, η_p^2 = .022. In conclusion, there is a significant difference in the learning gain on syntactic accuracy in the written description of places between DI and PF students.

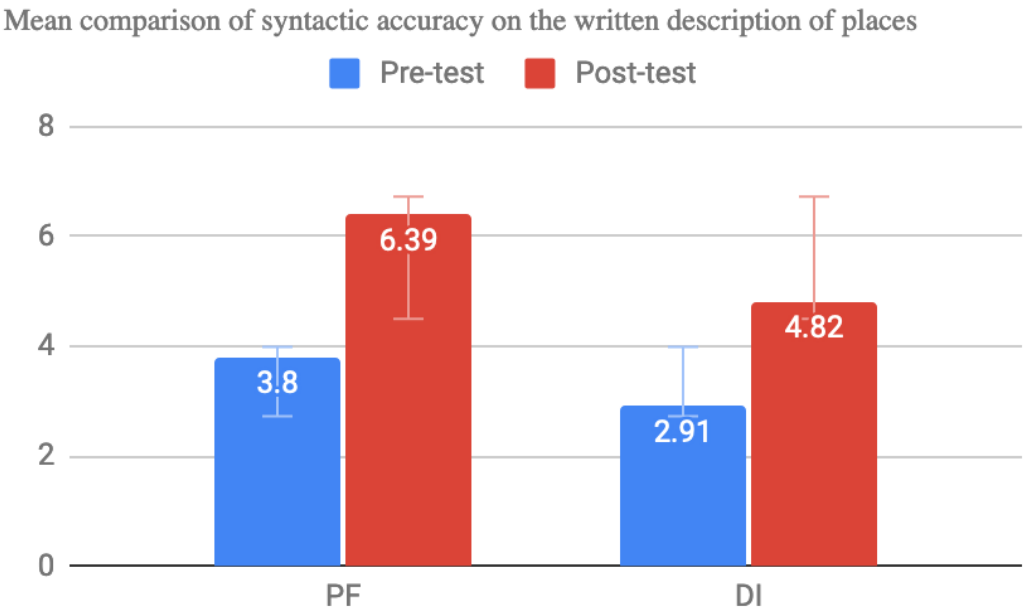


Figure 10 Learning gain on the syntactic accuracy of the written description of places

The result shows a significant effect of pedagogical strategies (favouring PF) on the syntactic accuracy of the written description of places. This suggested that the hypothesis, "there was no significant difference between PF and DI groups on the syntactic accuracy of the written description of places" was rejected at $p < .001$, one-tailed test. PF group outperformed DI group on the procedural knowledge (syntactic accuracy) of the written description of places (see Figure 10).

4.1.3. Students' learning gain in the declarative knowledge of the written description of places.

Both PF and DI groups started from the same prior knowledge of the lexical accuracy in the written description of places. An independent samples t-test reported no significant difference between PF and DI groups on the lexical accuracy pre-test score on the written description of places $t(110) = -.602$, $p = .548$, two-tailed. In addition, a one-way analysis of variance reported so significant effect of gender ($F(1, 110) = 1.771$, $p = .186$, two-tailed) and origin ($F(2, 109) = .056$, $p = .946$, two-tailed) on the lexical accuracy of the pre-test score on the written description of places. There was, however, a significant effect of institutions $F(2, 109) = 5.242$, $p = .007$ on the pre-test score of lexical accuracy of the written description of places. Thus, the effect of the institution was controlled as the covariate in the comparison between PF and DI groups.

A one-way analysis of co-variance reported no significant effect of pedagogical strategies on students' post-test score on the lexical accuracy in the written post-test between PF and DI groups, $F(1, 109) = 2.478$, $p = .059$, one-tailed, $\eta_p^2 = .022$. Neither was a significant effect of institution $F(1,109) = .810$, $p = .185$, one-tailed, $\eta_p^2 = .007$ on the students' post-test score on the lexical accuracy in the written post-test between PF and DI groups. In conclusion, there was no significant effect of pedagogical strategies on the students' learning gain on lexical accuracy on the written description of places.

The result shows that there was no significant effect of pedagogical strategies on the lexical accuracy of the written description of places. It suggested that the null hypothesis, "there was no significant difference between PF and DI groups on the lexical accuracy in the written learning gain" was retained in $p = .059$, one-tailed test. The PF group performed equally compared to the DI group on the learning gain on lexical accuracy in the written description of places (see Figure 11).

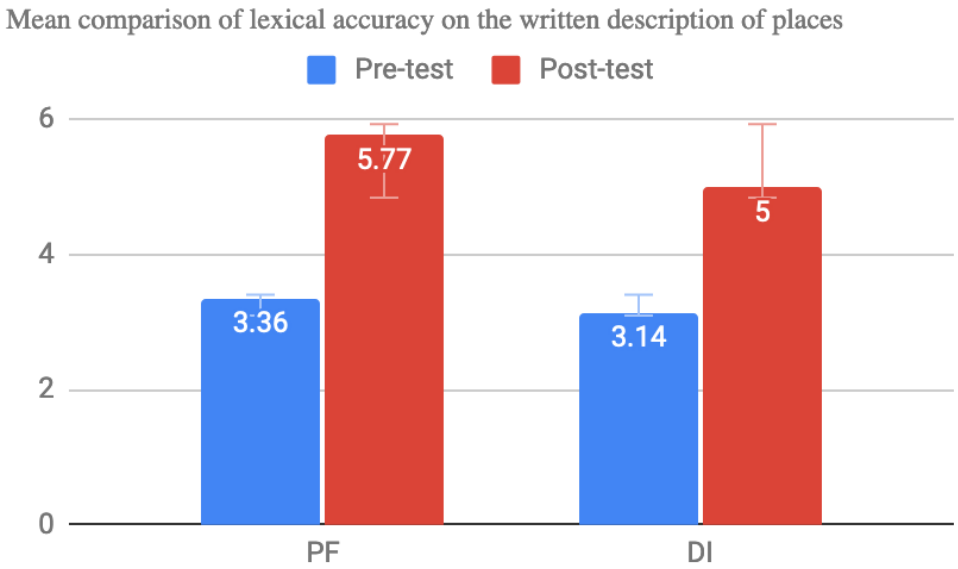


Figure 11 Learning gain on the lexical accuracy of the written description of places

4.2. Research Question 2: the Different Achievement of the Oral Description of Places

A summary of the spoken-based tests' results targeting the overall score, lexical accuracy, and syntactic accuracy scores in the spoken description of places is described in Table 6. Preliminary tests, including normality test, error variance, and independent observation, indicated that the ANOVA test could be applied to the post-test scores of procedural knowledge of describing places. The Shapiro Wilk's test reported that both DI and PF groups' scores were normally distributed (p-DI=.092, p-PF=.100). Levene's test based on mean, in addition, resulted in p =.478, which indicated the non-significant result. Moreover, the scores were obtained from independent tests. These results concluded that ANOVA could be applied to the post-test scores.

Table 6 Summary of the mean and standard deviation of the spoken test scores

	Lexical		Syntactic		Overall Spoken Test Scores	
	Experiment	Comparison	Experiment	Comparison	Experiment	Comparison
Pre-test	2.96(1.21)	3.20 (1.69)	2.77 (1.36)	2.79 (1.30)	5.73 (2.22)	5.98(2.64)
Post-test	5.14(1.86)	4.82 (1.85)	3.98 (1.79)	4.09 (1.83)	9.13 (3.33)	8.91 (3.12)

4.2.1. Learning gain on the overall oral test

Both PF and DI started from the same position in their spoken description of places. The independent-samples t-test reported no significant difference in the total score of pre-test on a spoken description of places between PF and DI students, t(110) = .543, p = 589, two-tailed. In addition, a one-way between-subject analysis of variance indicated no effects of gender (F(1, 110) = 3.371, p = .127, two-tailed); institutions (F(2, 109) = 1.172, p = .314, two-tailed); and origin (F(2, 109) = 1.400, p = .251, two-tailed) on the pre-test of oral description of places. We concluded that the pre-test scores of oral description places were equal between the PF and DI groups regardless the gender, origin, and institution.

A mixed between-within subjects analysis of variance reported no significant effects of pedagogical strategies on students' learning gain of the spoken description of places F(1, 110) = .124, p = .363, one-tailed, η_p^2 = .001 and no interaction between time and group F(1, 110) =.620, p = .217, one-tailed, η_p^2 = .006. We concluded that no significant difference was reported between PF and DI students on the learning gain of the spoken description of places.

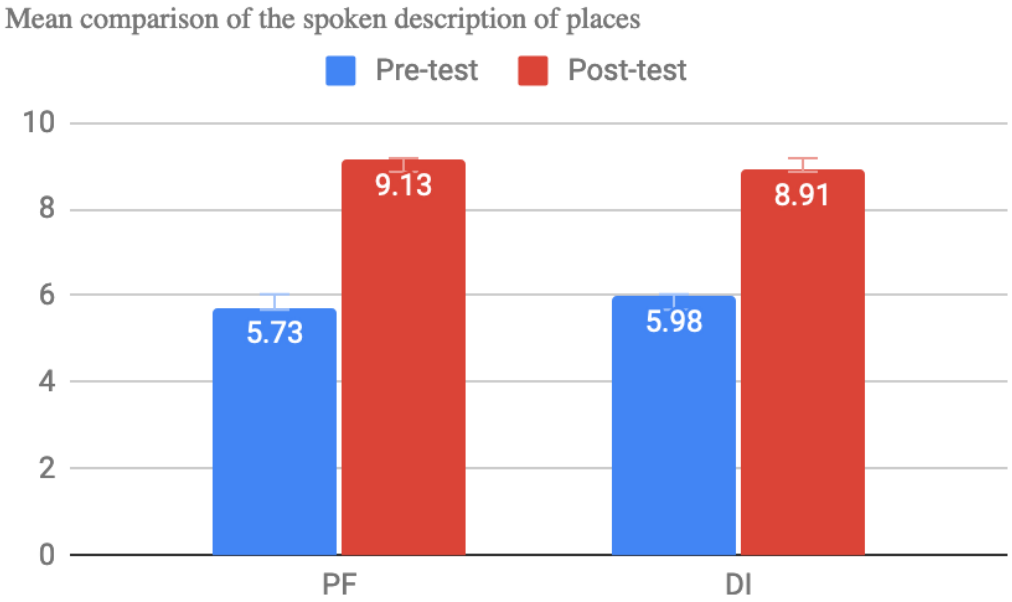


Figure 12 Learning gain in the spoken description of places

The result revealed that there was no significant effect of pedagogical strategies on the oral description of places. This indicated that the null hypothesis was retained in p = .363, one-tailed test. We concluded that PF students performed equally compared to DI students on their overall learning gain in the spoken description of places (Figure 12).

4.2.2. Students' learning gain in the procedural knowledge of the oral description of places.

The pre-test score of syntactic accuracy of the spoken description of places was equal between PF and DI. An independent-samples t-test reported no significant difference in students' pre-test scores of syntactic accuracy between PF and DI students t(110) = .071, p = 944, two-tailed. Neither did a one-way analysis of variance report a significant effect of gender F(1, 110) = .933, p = .336, two-tailed; institution F (2, 109) = 2.146, p = .122, two-tailed**;** and origin F (2, 109) = 1.985, p = .142, two-tailed on students' pre-test scores on the overall spoken description of places between PF and DI groups. Thus, we concluded that both PF and DI groups started from the same level of syntactic accuracy of the spoken description of places regardless of their gender, origin, and institution.

A mixed between-within subjects analysis of variance reported no significant effects of pedagogical strategies on the students' learning in the lexical accuracy in the spoken description of places F(1, 110) = .067, p = .399, one-tailed, η_p^2 = .001 and there was no interaction between time and group (pedagogical strategies) F(1, 110) = .063, p = .402, one-tailed, η_p^2 = .001. We concluded that there was no significant effect of pedagogical strategies on students' learning gains on the lexical accuracy in the spoken description of places.

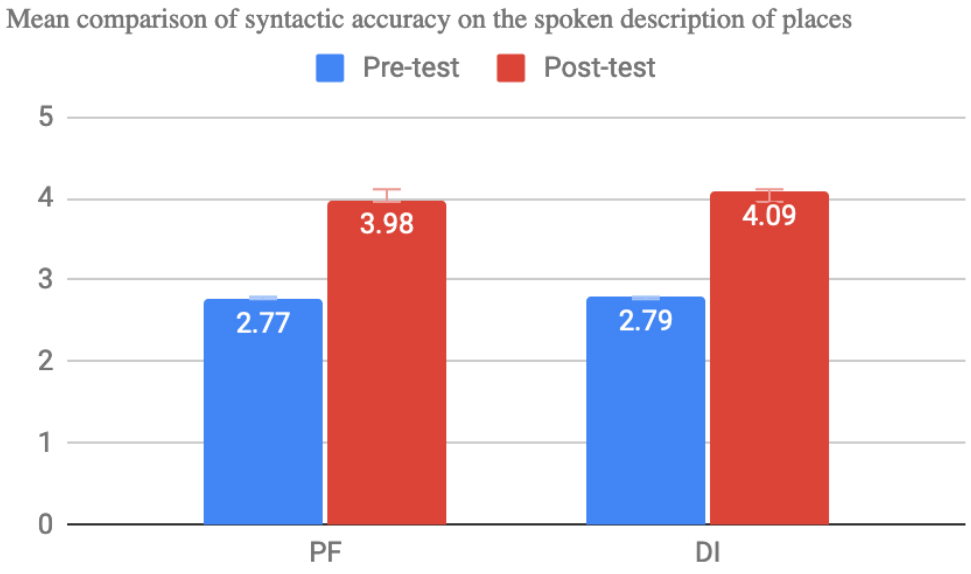


Figure 13 Learning gain in the syntactic accuracy of the spoken description of places

The result showed no significant effect of pedagogical strategies on the syntactic accuracy in the oral description of places, suggesting that the null hypothesis was retained in p = .399, one-tailed test. We concluded that students in the PF group performed equally

compared to those in DI group on their learning gain of the syntactic accuracy in the spoken description of places (see Figure 13).

4.2.3. Students' learning gain in the declarative knowledge of the oral description of places.

An independent samples t-test reported no significant difference between the PF and DI groups on their pre-test score of lexical accuracy t(110) = .837, p = .404, two-tailed on a spoken description of places. A one-way ANOVA, likewise, reported no effect of gender F(1, 110) = 2.373, p = .126; university F(2, 109) = 1.011, p = .367, two-tailed; and origin F(2, 109) =.836, p = .436, two-tailed. We concluded that PF and DI students' pre-test score on lexical accuracy was equal regardless of their gender, origin, and home university.

The result of the normality test showed that the data was not normally distributed. Therefore, a Mann-Whitney U test was carried out to test the hypothesis. A Mann-Whitney U test indicated that there were no significant differences in the learning gain between the PF (Mdn = 2.5, MR = 59.38) and DI (Mdn = 2.0, MR = 53.62) groups on the lexical accuracy of the oral description of places, U = 1.729, Z = .948, p =.172, one-tailed.

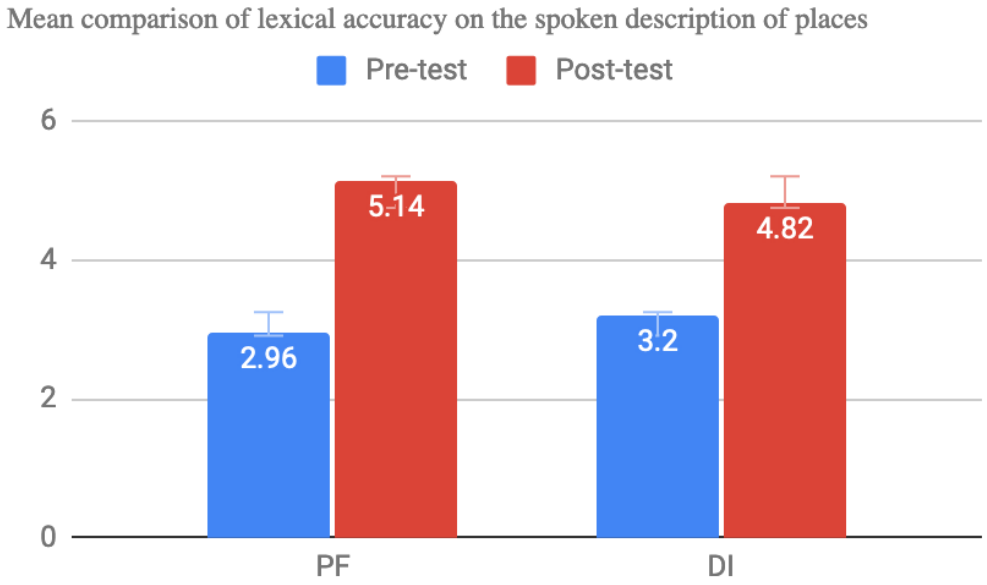


Figure 14 Learning gain in the lexical accuracy of the spoken description of places

The result showing that there was not a significant effect of pedagogical strategies on the lexical accuracy in the oral description of places retained the null hypothesis (p = .172, one-tailed test), "there was no significant difference between PF and DI groups on the learning gain of the lexical accuracy in the oral description of places." We concluded that students in the PF group performed equally to those in the DI group on their learning gain of the lexical accuracy in the oral description of places (see Figure 14).

4.3. Research Question 3: Student's Behaviour during the Instructional Video

Research question 3 dealt with the students' behaviour during the consolidation phase based on Greene and Azevedo's macro-level analysis of self-regulated learning (Greene & Azevedo, 2009). It was grounded from Winne and Hadwin's metacognition and self-regulated learning phases – task definition, goal setting and planning, study tactics, and reflection and evaluation (Winne, 2017a; Winne & Hadwin, 1998, 2008) – and Greene & Azevedo's (2009) task difficulty and demands. The project, however, focused on monitoring and study tactics (learning strategies) – provided they are essential characteristics of highly self-regulated learners (Pintrich, 2000; Winne & Perry, 2005; Zimmerman, 2005). Another reason was that the focus of the observation was during the consolidation, in which students watched the instructional video. Students' responses to the task difficulty and demand were also observed given that they were essential parts of the consolidation phase.

The video observation data revealed that participants deployed four macro-levels of self-regulated learning, namely motivation, cognitive learning strategies, task difficulty and demand, and monitoring (metacognitive) learning strategies (Table 7). Quantitative and qualitative analysis was presented in the following sections.

Table 7 Summary of Macro-level Self-regulated Learning

Self-regulated learning skills	Frequency		Mean	
	PF	DI	PF	DI
Motivation +	45	18	0.80	0.32
Motivation -	11	38	0.20	0.68
Cognitive Learning Strategies	125	28	2.59	0.54
Monitoring Learning Strategies	68	3	0.86	0.02
Task difficulty and demands	8	5	0.14	0.09

4.3.1. Motivation

As seen in Table 7, motivation (both + and -) became the second mostly enacted behaviours during explicit instruction in the consolidation phase. The students' motivation in watching the instructional video was shown in their persistence, attention, and interaction with the instructional video in the consolidation phase. Students who finished watching the video attentively and responded to the teachers were considered to show motivation (+). In contrast, participants who did not finish watching the video or watched the video while doing something else was considered to enact motivation (-). When students showed the motivation (+) behaviour, it was hypothesised that they valued the task and had goal orientation toward the video (Bong, 2001; 2004).

Table 7 shows that The PF group (45) deployed more motivational behaviour while watching the instructional video than the DI group (18). Independent samples t-test reported that PF students showed significantly more motivational behaviour t(110)= -6.808, $p< .01$. They responded to the teacher, mimicked, and practised the target expressions, reacted to new terms, and watched the video interactively.

Conversely, the DI group showed significantly higher motivation (-) behaviours compared to the PF group during the instructional video t(110) = 6.879, $p< .01$. Watching the video straight from the beginning till the end with no interactions was the most-observed behaviours among DI students. Seven DI students did not look at the screen when playing the instructional video or doing something else while watching. In conclusion, students in the PF group had stronger motivation to watch the instructional video based on their observed behaviours. Table 8 shows motivation (-) behaviour from the observation data.

Table 8 Frequency of Evidence of Motivation – between PF and DI

Behaviour	Frequencies		Percentage	
	PF	DI	PF	DI
Watched only	1	14	1.96	27.45
No continue	6	9	11.76	17.65
No screen contact/ sleepy	1	7	1.96	13.73
Talking to other people	0	5	0.00	9.80
Jumping windows and scenes	1	4	1.96	7.84
Playing pen	0	3	0.00	5.88
TOTAL	9	42	17.65	82.35

The following paragraphs discuss the commonly observed motivation (-) behaviours during the consolidation phase. The first commonly observed behaviour (27.45%) among DI students was watching through the video. They were observed to play the video passively. No pointer, mouth, and body movements were observed during the 8-minute video. One example was seen in the observation notes of DI21. "She watched the video without doing anything. No lips' moves, no body-moves, and eyes straight to the screen." (Obs-Mot(-)-DI21)

The second commonly observed behaviour showing motivation (-) action was incomplete access to the instructional video. Despite the statistically equal occurrence, t(101.96)= -6.808, $p < .001$ between PF and DI students, there was a difference in the way they watched the instructional video. The majority of DI students watched the parts passively while most PF students (5 out of 6) watched the video actively (for an example, see the

observation notes). The observation notes show that DI24 played the video without positive responses while PF18 responded to the instructional video actively despite the partial access.

"She only listened to the recording. No lips' moves, no body-moves, and eyes straight to the screen. Did not come back after the quiz."	Obs-Mot(-)-DI24
"He tried to answer the aspects of describing places. He paused the video on aspects of describing places, took notes, He mimicked/ read the vocabularies introduced in the video. He took notes on the stilt house. He repeated the unknown prepositions. He did not come back to the video after the quiz."	Obs-Mot(-)-PF18

The other interesting behaviours from the majority of DI students were playing the video while doing something else or without seeing the screen. It was coded as motivation (-) as attention was not paid on the video despite the play. A small number of students played the instructional video without eye contacts because they were sleepy (Obs-Mot(-)-DI57), talked to other people (Obs-DI110), or jumped windows while playing (see Obs- Mot(-)-DI13).

"She played the video until the finish, but her eyes were off-screen."	Obs- Mot(-)-DI28
"He played the video. He was sleepy."	Obs- Mot(-)-DI53
"She played the video jumping from one scene to the other. And she kept moving the windows."	Obs- Mot(-)-DI10
"She watched the video while moving the windows."	Obs- Mot(-)-DI13

In sum, there was a difference in motivation to watch the instructional video between DI and PF, with significantly higher motivation for PF students. The difference was observed from their internal factor like being sleepy and their willingness to respond to the video and external factor like the temptation to communicate with other people while watching the instructional video. Another interesting finding from the observation is the learning strategies students used while watching the instructional video.

4.3.2. Cognitive learning strategies.

The observational data revealed that PF group (81.70%) dominated the employment of the learning strategies compared to DI group (18.30%), in which the difference was statistically significant ($t(110) = -6.197$, $p<0.01$). The primary learning strategies (more than ten occurrences) included memorising, practising, being interactive to the instructional video, reviewing, and taking notes. At the same time, DI enacted "practising the new expressions" (12 occurrences) as the primary learning strategies followed by memorising vocabularies. Table 9 shows the summary of the learning strategies from both groups.

Table 9 Summary of the evidence of learning strategies

Learning strategies	PF	(%)	DI (%)	(%)
Memorisation	34	22.22	8	5.23
Practices	33	21.57	12	7.84
Interaction*	16	10.46	2	1.31
Reviewing	14	9.15	0	0.00
Taking notes (copying texts from the recording)	13	8.50	0	0.00
Re-reading (revisiting a section)	8	5.23	1	0.65
TOTAL	125	81.70	28	18.30

On average, a participant in the PF group executed 2.6 learning strategies during the instructional video. However, one participant could use more than eight strategies during the instructional video. For example, a participant (PF16) was observed to carry out eight learning strategies during the consolidation period (10.27 minutes). She started with positive motivation at the beginning of the video after describing the place followed by five kinds of cognitive learning strategies and two monitoring strategies.

Obs-Complete-PF16	Codes
"… Eyes on the screen.	Mot+
Happily watched the video.	Mot+
Often said "Oh!"	MON-JOL-
Paused after "melodic ocean waves".	LS-Ps
… took notes.	LS-TN
Paused and Notes taken on the aspects of describing places.	LS-TN
She read the vocabularies introduced in the video.	LS-Pr
Repeated the vocab in the video.	LS-Rw
Said, "Ohh…! when "shade" was introduced in the video.	MON-JOL-
She did the N-adj combination question.	LS-Pr
Paused and read the formula…"	LS-Ps and LS-Pr
TOTAL: (2Mot+, 2MON-JOL, 2LS-Ps, 2LS-Tn,1LS-Rw, 3LS-Pr)	

Notes: Mot (motivation), NON-JOL (monitoring, justification of learning), LS-Ps (learning strategy – pausing), LS-TN(learning strategy – taking notes), LS-Rw (learning strategy – reviewing)

On the other hand, one participant was observed to enact no learning strategies than silently watched the instructional video in the consolidation period. They played through the video without doing anything. An observation report of DI21 indicated that the participant enacted no learning strategies.

Obs-Complete-DI21	Codes
"she watched the video straight from the beginning to the end without doing anything."	Mot (-)
TOTAL: 1Mot(-)	

In conclusion, PF students used significantly higher self-regulated learning in terms of learning strategies. It was evidenced that both groups had the same main strategies, i.e., memorisation, practice, and active interaction with the instructional video. However, the number of learning strategies employed by PF students was significantly greater than by DI students. It is interesting to see if this trend also occurred in a vital aspect of SRL, monitoring.

4.3.3. Monitoring learning strategies

The learning monitoring was executed by comparing students' description in the first task to the description model presented at the onset of the consolidation phase. The strategies comprised negative Judgement of Learning (JoL-) and positive Feeling of Knowing (FoK+). JoL- was coded when students realised the errors they made in the previous task. It commonly happened during the model description or the instruction. FoK+ happened when students realised that they just knew a new knowledge. It took place during the explicit instruction in the instructional video.

Similar to the previous trends, a significantly higher use of monitoring strategies among PF students than DI students in the consolidation phase was reported. The observed monitoring strategies during the instruction included JOL- and FOK+. An independent sample t-test reported a significant difference in the use of negative JOL ($t(110) = -4.227$, $p<.01$), favouring PF group. Likewise, a significant difference was reported on the use of positive FOK ($t(110) = -3.798$, $p < .01$) denoting greater use among PF students compared to DI students.

Table 10 Evidence of FOK+ and JOL-

Observed behaviours	FOK +	Observed behaviours	JOL-
Nodded head	9	Paused the video on certain parts	20 (2)
Say the word in a different intonation	5	Um…., Oh…	10
Oh! Ah! Aha!	3	Nodding	8
Repeated the word, pause	3	Repeat the word	4
Facial expression	1	Pause	1
statement	1	Confirmation by yelling	1
	22		44

Students expressed their FOK+ and JOL- differently. Nodding head was the most frequently observed behaviour to show FOK+, while the verbal statement was the least. An interesting marker of FOK+ was the " Aha!" statements which can be in the forms of saying the target expression loudly, happily repeating the expressions in high intonation (emphasis),

or facial expressions. The JOL- was mostly observed in the students' weak "ums" or "oh". Table 10 shows the complete lists of evidence of FOK+ and JOL-.

The use of JOL- was often observed to come together with the use of FOK+. The majority of PF students often confirmed to have used wrong expressions in the first task and felt to have learned new expressions of *"Now I know something new!"* at the same time. One example was observed from PF10 during the instructional video and from his exit survey. He was observed to confirm that he used the wrong vocabularies (JOL-) and his facial expression denoted that he felt to have learned new expressions (FOK +) at the same time. In the exit survey, he mentioned that he picked up vocabularies from the video.

Obs-Complete-PF10	Codes
"He often yielded confirming that he has used wrong vocabularies in the yellow ball	JOL(-)
He smiles for new vocabularies or expressions	FOK(+)
he changed her face (note: excited) when he found out new expressions	FOK(+)
TOTAL: 1JOL(-), 2FOK(+)	

Exit survey – PF10	Codes
"He mentioned that he picked up some vocabularies from the video."	FOK(+)
TOTAL: 1FOK(+)	

However, feeling of knowing (FOK+) was sometimes observed without the occurrence of JOL- or vice versa. One PF student (PF06) was coded to confirm that she understood new expressions by saying the word loudly or softly in a murmur. These reactions were believed to be expressions of knowing new knowledge from the instructional video. In this particular student, she used FOK+ as the monitoring strategy for her learning, yet she did not use JOL- like the previous example. PF 02 was also observed to execute one monitoring strategy (JOL-), when she said: **"oh, stilt house!"** This response was translated into "So, it was stilt house" which infers that she realised that she did not describe "stilt house" correctly during the idea generation and exploration (JOL-). These two examples denote that the occurrence or JOL- was not always followed by FOK+ based on the students' observed behaviours. For example, a student realised that they had made mistakes during the exploration and generation task by mimicking the target expression in high voice (see Obs-PF6).

"She murmured "oh stilt house!", then she also murmured "shade"… She said the word "stilt" very loudly ..."	FK+-PF06
"…oh, stilt house!" and paused and took note while kept saying the word. …	FK+-PF02
"Oh, stilt house?" She said this when the video modelled the description.	FK+-PF2
"Ah…" "Erm…" She murmured when new vocabularies (straw shelter, blissful, on top of, bristle-like, stilt house) were used in the model or explained during the instructional video.	PK+-PF3
Students said the words loudly while watching the instructional video.	FK+-PF6

4.3.4. Task difficulty and demands

Reactions to task difficulties and demands were observed both in PF and DI groups by turning on sub-titles and asking for help from the conversation companion. The same number of students in both groups turned on sub-titles, but only students from the PF group asked for help from the conversation companion as their help-seeking behaviour during the consolidation phase. All DI students turned on the sub-title from the beginning of the instructional video, yet three PF students turned on the sub-title after listening to sections of the description model. It was predicted that the PF students turned on the sub-title as a reaction to a monitoring process (FoK-). In contrast, the DI students turned on the sub-title as a practice that they might have done before the treatment. Therefore, the decision was not taken due to a process of monitoring.

In terms of asking for help, PF students attempted to use the strategy when they experienced difficulties (Table 11). However, they did not get the answer as there was a restriction of conversation companion's role. They could not answer content-related questions during their work.

Table 11 Students' reaction to the task demands and difficulty

Task Difficulties and Demands	PF	DI
Help-seeking behaviour		
- Turning on subtitles	5	5
- Asked for help from the conversation companion	3	0
TOTAL	8	5

In summary, the observational data showed that PF students enacted more self-regulatory skills during the instructional video in the consolidation phase. They watched the video in higher motivation, as shown in their reaction to the video. They enacted more monitoring strategies toward their learning and more learning strategies while watching the video. Unlike PF students, DI students were reported to watch the video passively (unmotivated) as indicated in the number of within video drops and less monitoring and

learning strategies while watching the video. Therefore, we can conclude that PF and DI students reacted differently toward the instructional video. It can be the results of the different intervention they got.

4.4. Summary

This research compared the achievement between PF and DI in the written achievement in the overall written description, the vocabulary use in the written description, and the syntax use in the written description of places. In the overall written description, the PF group showed a significantly higher learning gain than the DI group. Likewise, PF students outperformed DI students in the use of syntax in the written description of places. In the lexical accuracy, however, both groups attained equally substantial learning gains on the vocabularies used in the written description of places.

Unlike their performance in the written description, PF and DI groups performed equally in all aspects of the comparison in their achievement in the spoken description of places. Both PF and DI groups gained equally substantial learning gain on the verbal description of places. The equal gain was also found in the lexical and syntactic accuracy in the verbal description of places. Both groups gained equally in the use of vocabularies and syntax in the spoken description of places.

The evidence from the students' behaviour during the consolidation phase revealed that the PF group used significantly more self-regulating behaviours compared to DI group. PF Students utilised significantly greater monitoring and learning strategy skills during the consolidation phase. They employed judgements of learning (JOL-) and feelings of knowing (FOK+) substantially more frequently than DI students. The evidence from the video also suggested that PF students enacted significantly more positive learning strategies, such as memorisation, practice ("draw" in Greene & Azevedo, 2009), reviews, note-taking, revisiting sections (this equals to "re-reading" in Greene & Azevedo, 2009), and interaction. In addition, turning on sub-titles and asking for help from others have been their strategies to react to the task demands and task difficulty.

The additional finding showed that PF students watched the instructional video with significantly higher motivation than DI students. They showed more enthusiasm and responded more actively to the video. They kept their attention to the video, responded to the teacher's questions, practised the target expressions as expected. On the other hand, the DI students played the instructional video, but paid no attention to it, for example, without eye contacts to the video. They were also observed to talk to other users, jump between windows, or look for sections while playing the instructional video in the consolidation phase. This

evidence supported that PF students showed more motivation to watch the instructional video in the consolidation.

CHAPTER FIVE: DISCUSSION

This study aimed to compare a Productive Failure (PF) instructional design to a Direct Instruction (DI) instructional design that involved a virtual learning environment named Second Life. The project was carried out in English as a foreign language learning, in which PF has not been carried out before. The limitation was set to the lexical and syntactic accuracy of the written and spoken description of places. Findings of the study suggested that: (a) PF students improved significantly higher in the written description of places with a significant difference in the learning gain of procedural knowledge—favouring PF students—yet an equal learning gain on the declarative knowledge; (b) PF students improved equally compared to DI students in the spoken description of places, both in declarative and procedural knowledge; and (c) significant differences were reported between PF and DI groups on their self-regulated learning strategies during the consolidation phase based on the qualitative data from the video captures. The discussion first considers how and why the findings were similar to (or different from) our expectations. Second, the theoretical and practical implications were discussed.

5.1. Research Question One: Learning Outcome in the Written Description of Places

Research question one was concerned with whether there were different learning outcomes in the written assessment, especially on declarative (lexical accuracy) and procedural knowledge (syntactic accuracy) between PF and DI. As expected, PF students outperformed DI students on the procedural knowledge in a medium effect size but performed equally on declarative knowledge. It indicated that PF students were more capable of not only deciding vocabularies to complete sentences but also applying the appropriate syntax for the words to make grammatically correct sentences. Both groups performed equally on the declarative knowledge, a task that requires students to recall vocabularies by looking at either still pictures or pictures in a video.

The effect size of the significant difference in the procedural knowledge ($\eta p^2 = .088$) was in the medium to the medium-high range (Cohen, 1977), which is lower compared to the previous study in Lai et al. (2018) with $\eta^2 = .158$. The reason might be that the intervention was shorter (80 – 90 minutes per person) than other previous PF-related studies. Students described a virtual beach named Virtlantis, accessed an eight-minute consolidation video, and described another virtually outdoor place named NCI beach. It was a one-day participation. Longer duration of studies using the same experimental treatment could determine if a more significant effect size could result, suggesting more effective language learning.

In this study, students learned a new vocabulary from a picture or a video section (declarative knowledge) and the appropriate syntax of the vocabulary in noun or adjective phrases and the right syntax to make correct sentences to describe a place (procedural knowledge). For example, they learned the word "bristle-like" and "plant" from high-lighted pictures in a video, in which both PF and DI groups had an equal result. For a procedural knowledge assessment, they used the noun phrase of "a bristle-like plant" in a sentence of "I can see a bristle-like plant on top of the black rock." PF group performed significantly better in procedural knowledge, in which they correctly applied the syntax in the written sentences to describe a place.

The result was consistent with the previous studies comparing PF and DI in Maths (Kapur, 2015; Loehr et al., 2014), statistics (Jacobson et al., 2017; Kapur, 2011, 2012; Kapur & Bielaczyc, 2012), and Nano-science (Lai et al., 2018). Those studies reported that PF students outperformed the DI students in procedural knowledge but performed equally on declarative one. Procedural knowledge is the ability to correctly apply the learned knowledge (Rittle-Johnson & Schneider, 2015) while declarative knowledge is the knowledge of facts, concepts, and procedures (Ullman, 2001). In this research, declarative knowledge refers to the accuracy of vocabulary use, which is the relation between meaning and form, while procedural knowledge refers to the appropriate use of vocabularies in syntactically correct sentences (Paradis, 2009).

While solving complex problems was typical for the first challenge in most PF studies (for more detail on PF design, see Kapur & Bielaczyc, 2012), the current project used a communicative task as the first challenge. The majority of participants (N=15) in the pilot study reported that the task was complex relative to their English proficiency, especially on the vocabulary, pronunciation, and syntax. This study focused on the lexical and syntactic components of describing places, which have been Indonesian English learners' challenges (Cahyono, 2008). In addition, they are measurable within a short period of treatment (Saito, 2014; Yoshii & Flaitz, 2002). It was also complex based on The Triadic Componential Framework for Task Classification (Robinson & Gilabert, 2007). However, video observation indicated that they were not frustrated as they have been familiar with a number of basic vocabularies and syntax used in the description.

5.2. Research Question 2: Learning Outcome in the Oral Description

The results related to RQ2—the differences between PF and DI in declarative and procedural knowledge in the spoken description—were not in line with our hypothesis. There were no significant differences between the PF and DI groups on the spoken description of

places in both declarative and procedural knowledge. There might be a different mechanism between written and spoken production in L2 language learning. It may be that neither the PF nor the DI conditions provided the students with sufficient time or opportunities to use the new target language syntax into spoken production of the language (Ahmadian & Tavakoli, 2011; Finardi, 2008; Lambert, Kormos, & Minn, 2017).

Students practised the target syntax twice (in PF) and three times (in DI) after the subsequent instruction (post-test included) in this research. Previous studies on oral production suggest at least two repeat tasks for fluency purposes (Bygate, 1999; Finardi, 2008) and at least four similar tasks repeated in a more extended period for accuracy purposes (Lambert et al., 2017). Based on Levelt's (1999) model of oral production, students undergo three processes of oral production, i.e., conceptualisation, formulation, and articulation. It is challenging for foreign language learners (especially intermediate) to control the three processes under time pressure, especially during the speech act's real-time processing. It has been commonly understood that due to the real-time processing of speech act, resulting in more fragmented grammar. It is challenging for foreign language learners to focus on accuracy, fluency, and complexity simultaneously. Those processes need more extended time and more practices for students to automatically use target expressions orally.

The result on the procedural fluency in the oral production is partially inconsistent with Loibl and Rummel's (2014a, 2014b) studies suggesting that it was the number of practices after the instruction that affected procedural fluency. The more practices after instruction, the better the improvement of procedural fluency. However, this research denoted that no significant difference was reported between PF and DI on syntactic (procedural fluency) of the spoken description of places. Descriptively, there was only .09 difference favouring DI group, which can be ignored. Therefore, more studies need to be performed to test Loibl and Rummel's stance on procedural knowledge development. The next discussion covers the self-regulated learning analysis on the students' behaviour during the consolidation.

5.3. Research Question 3: Self-regulated Behaviour during the Consolidation Phase

The results related to research question 3, the self-regulated learning behaviours during the consolidation phase, was consistent with our expectation. There were three main findings from the qualitative data. First, PF students used significantly more monitoring strategies, the judgement of learning (JOL) and feeling of knowing (FOK). The description model in the consolidation video made the monitoring possible. They compared their description to the canonical one in the consolidation while watching the description model

and listening to the instruction. One of PF students answered “I'm learning, and I compared with my way to describe in the yellow ball” to a question in the exit survey.

Second, PF students used significantly more cognitive strategies compared to DI students. They practised the target vocabulary and expressions, took notes, reviewed video segments in pauses and playbacks, and sought help. The last finding was that PF students showed significantly more motivational behaviours while watching the video than DI students. They showed enthusiasm by watching the instructional video from the beginning until the end. They showed positive engagement, e.g., by responding to the teacher, taking notes, and expressing success and failures in their learning. In contrast, DI students did not show their positive attention when watching the instructional video. The majority of DI students watched the instructional video while doing something else, without concentration, or jumping windows.

Contextualised into the previous studies on PF, there were no studies on SRL in PF-based instruction to date, to the best of our knowledge. However, the finding of this research on the superiority of PF students in self-regulated learning has been predicted. It corroborates previous studies on SRL suggesting the relationship between prior knowledge and self-regulated learning skills (Moos & Azevedo, 2008a; Taub & Azevedo, 2017; Taub et al., 2014; Trevors et al., 2014). These studies reported that students with high prior knowledge deployed more effective cognitive strategies and higher metacognitive strategies than those with lower prior knowledge. The current studies found that students in the PF group, who were presumed to have activated more prior knowledge, deployed more cognitive and metacognitive strategies. They performed more meta-cognitive monitoring strategies, such as JOL and FOK, and the cognitive learning strategies compared to DI students who did not activate their prior domain knowledge.

Additionally, the finding that PF students had higher motivation than DI students corroborates PF-related qualitative studies on the classroom dynamics (Jacobson et al., 2017; Pathak et al., 2011) that PF students are more motivated in the consolidation phase. They showed deeper engagement with the learning model or higher curiosity with the instruction (Pathak et al., 2011) and asked more questions and solved problems more collaboratively (Jacobson et al., 2017a). While those previous studies reported classroom-based instruction, the current project reported video-based instruction in that the interaction was reported between the students and the instructional video. PF students responded actively to the questions or instructions from the virtual teacher in the instructional video.

5.4. Theoretical Explanation

Previous PF-related studies suggested that PF's cognitive mechanisms played a crucial role in its superiority over DI on procedural fluency (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015). Provided they have activated and differentiated their prior knowledge, acknowledged their knowledge gaps, and better abstract the schema, PF students learned and gained better in procedural knowledge. The discussion of the results is based on PF's mechanisms and self-regulated learning theorising.

5.4.1. Prior knowledge activation and differentiation

Prior knowledge activation and differentiation is the theoretical explanation of the PF's superiority on the procedural fluency. Finishing the communicative task in the idea generation and exploration phase, PF students attempted to linguistically respond to the video task, allowing them to activate their prior knowledge. However, they responded in a way that was only partially accurate provided they have incomplete prior knowledge about English, and they spent a more extended time yet produced fewer words in the first description task.

For example, students in the PF group knew the word "house" and the article "a/an" before the class began, as evidenced by their responses to the first task. Majority of them mentioned "a wooden house" or "a high house" with a modifier "above the sea" or "above the water." To successfully describe an elevated house, they need an adjective "stilt", which they have not successfully produced in the first task—which is an example of a knowledge gap. This recognised gap was then filled when they watched the description model, which was then enforced by the teacher's examples in the instructional video and when they did the quiz. Several PF students used the expressions correctly in the second similar task.

Students in DI group, in contrast, activated less prior knowledge as they did the first task after the instruction. Their equal overall pre-test score compared to the PF group indicates the same level of prior knowledge activation during the pre-test. Besides, despite the different pre-test for the procedural knowledge in written measurement, the pre-test was carried out two weeks prior to the treatment. Therefore, we can ignore the pre-test as a tool to activate prior knowledge between the two groups. Theoretically, they could have completed the first task better due to the preceding instruction, but they produced statistically equal words to PF students in the first challenge. DI students might not do the first communicative task seriously provided that they passively watched the instructional video in the consolidation phase. They might experience "false mastery", feeling that they have mastered the content of the instruction so that they do not see the benefits of watching seriously. It is

also possible that as they did not acknowledge their knowledge gaps correctly, they could not focus on the sections they needed to learn.

Students' prior knowledge was not only activated but also differentiated in the first task. They could describe the place only using their prior knowledge provided that they did not get any helps either from the conversation companion (more knowledgeable others) or from other online resources. During the consolidation, students in PF had the opportunity to compare their initial description to the model, which then followed the targeted concepts explained during the instruction, which supports prior knowledge differentiation. Students in the PF group could confirm which prior knowledge was suitable for the instruction's target concept. For example, a number of PF students activated their prior knowledge of "on" "above" or "on the top" to refer to a particular position of something sticking above something else. Listening to the canonical description in the consolidation phase, PF students knew "on top of" should have been used instead of "on", "above", or "on the top". This particular skill is useful for foreign language learners to use contextually suitable expressions. Prior knowledge activation and differentiation may help knowledge gap recognition, which is the focus of the next discussion.

5.4.2. Knowledge gap recognition

Comparing and contrasting students' solution to the canonical solution enables knowledge gap recognition (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), which makes students better able to discern and understand the concept taught in the consolidation phase (Kapur, 2011). The observation data and the exit survey found that the majority of PF students acknowledged their knowledge gaps of the target expressions in describing places. Their descriptions of the first task indicated that they had gaps in specific vocabularies and syntax used to describe places. One example that most PF students missed in their first description was the word "a straw shelter", which they described as "an umbrella". They knew the word "umbrella" for "shelter" and no word for "straw." They noticed the gaps of "straw" and "shelter" only after they compared their description to the model in the consolidation phase. in addition, answering an exit survey question about their feeling and activities when listening to the description model in the consolidation phase, one of PF students answered: *"Confused because I don't know what I should describe in my speaking, in the other side I haven't many vocabularies to speak in that video."* The answer indicated that he experienced an impasse "confused" and aware about the lack of vocabularies to do the first description.

On the contrary, most DI students missed the discussion on the use of the appropriate word for "shelter" from the instructional video, as indicated in the absence of this word in the majority of DI's second description. Instead, they still used "umbrella" instead of "shelter." It may be that DI students might have experienced "false mastery", feeling that they think they knew something that they did not know, which affected their understanding of the target concept and their attention to the instruction. In addition to prior knowledge activation and knowledge gap recognition (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), the schema abstraction during the comparing and contrasting activity plays a prominent role in the PF students' superiority in the procedural knowledge (Jacobson et al., 2020).

5.4.3. Schema abstraction

Schema abstraction, or repairing mental model in Chi (2000), is a knowledge construction process connecting prior knowledge and just-learned knowledge. As prior knowledge (old knowledge) activation and knowledge gap recognition are facilitated in PF, schema abstraction should be more successful. The knowledge gap recognition helps to sort which old information is needed to build new schemas (Jacobson et al., 2020). Completing the knowledge gaps during the instructional video can be associated with finding critical pieces of an unfinished puzzle, the targeted schema. The gaps, the missing parts, need to be discovered in the consolidation phase when students compared and contrasted their description to the canonical description. The targeted schemas might be different between one student to the others as they might have different prior knowledge. In Jacobson et al. (2020), for example, comparing and contrasting students' solution to the canonical solution enabled PF students to build a stable schema by adding the new pieces of knowledge into the existing one, called schema abstraction. To create a complete schema of how to describe a place, students had to realise their current schema, which is called prior knowledge, by describing places before instruction. Figure 15 illustrates the schema abstraction process experienced by the two groups.

In this study, students in PF groups benefitted from the description model and the activity of comparing and contrasting their description to the model, from which they acknowledged their knowledge gaps. In contrast, DI students' "false mastery" hindered them from the benefits of knowledge gap recognition as they may lose their focus in the instruction phase.

PF students in this study recalled old schema in the idea exploration and generation. They found which slots (gap) to fill with the new information as the results of comparing and

contrasting the consolidation process. The explicit instruction and the practices in the consolidation phase glued the information into the correct slots. As students have different old schema as prior knowledge, they have additional information to complete the old schema. When they are fixed, new schemas were changed, which differed as the students' progressed.

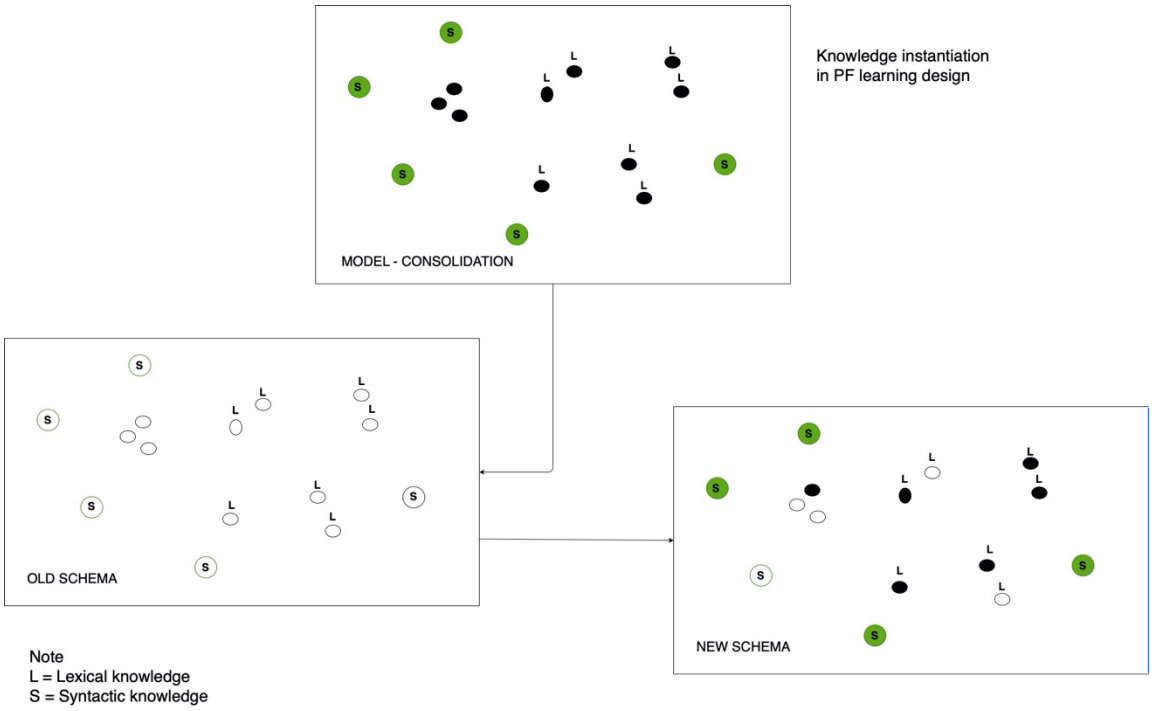


Figure 15 Schema abstraction among PF students

5.4.4. Self-regulated learning

Self-regulated learning skills might be a moderating factor of the PF students' success in learning. The task in the idea generation and exploration phase helps activate and differentiate prior knowledge (Kapur & Bielaczyc, 2012), and prior knowledge allows students to formulate sub-goals and plan their learning strategies in the real learning process (Pintrich, 2000; Winne, 2017a; Winne & Hadwin, 1998). The idea generation and exploration in PF design enable not only the prior knowledge activation but also differentiation. PF students could activate more prior knowledge from the authentic task in the idea generation and exploration so that they have more possibility to set accurate sub-goals and plan the learning (Moos & Azevedo, 2008b; Taub & Azevedo, 2017; Taub et al., 2014). Their motivation to access the instructional video and their attention to specific instructional video sections proved that they had set individual sub-goals and planned their learning strategies accurately. Unfortunately, the main data on the goal setting and strategy planning were not available, given that the observation was during the consolidation. In the exit survey, however, students claimed to plan their second task but not the first. For example, PF11 said

that she thought about how to describe the place and prepared the sentences needed in the description. Thus, prior knowledge activation and differentiation in PF might help accurate sub-goal setting and planning.

Comparing and contrasting students' task performance to the model (canonical) description helps PF students recognise their knowledge gaps (Jacobson et al., 2020; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014a, 2015), which can help PF students focus their attention to specific details during the instruction. In SRL theorising, this process enables them to monitor their learning (Pintrich, 2004; Winne, 2017a; Winne & Hadwin, 1998), which will affect the choice and evaluation of cognitive learning strategy use. The observation data showed that PF students used different cognitive learning strategies due to their recognised knowledge gaps and the judgements of learning resulted in the task execution from the previous process. For example, most PF students chose to practice by imitating certain target expressions instead of taking notes. The exit survey data showed that they had acknowledged gaps in their pronunciation of the target expressions. They were also aware of the speaking task after the instructional video so that imitating was their study tactic. Another interesting strategy for new vocabularies employed by PF students was revisiting a section of the instructional video several times on the word "shelter" as they acknowledged a gap at this specific vocabulary. Having known their learning gaps affected what learning strategy to pick.

As every PF student might have different learning sub-goals based on their knowledge gaps, they should apply more effective learning strategies during the consolidation. Like in the previous studies, PF students engaged deeper with the learning model (Pathak et al., 2011) and reacted to task difficulty by asking more questions (Jacobson et al., 2017a) compared to DI students. PF students in this study could specify sections to concentrate and revisit for more effective learning; accessed the video shorter (PF=12:32; DI=14:62) yet executed more positive study tactics; and showed more positive behaviour compared to the DI students. The exit survey revealed perceived knowledge gap recognition after the first task when the model description was introduced, and more learning strategies were observed among PF students compared to DI. Some behavioural expressions were found, such as "I see", "Oh!", nods, facial expressions, and loud word echoing after new vocabularies mentioned.

DI students, on the other hand, seemed to have activated less, if not no, prior knowledge. The observation data showed that most DI students watched the instructional video straight from the beginning to the end, and they did not pause or review sections of the

video. Therefore, prior knowledge activation and knowledge gap recognition of DI students were not as effective as those of PF students, and it might result in inaccurate (sub)goal setting. For example, some DI students used the same expressions in all description tasks. One more specific example is using the same and often inappropriate prepositions of places in all descriptions. A number of DI students used "on", "behind", "in front of" in their descriptions despite the discussion of new prepositions like "on the right side", "on the left side", or "on top of." The possible explanation is that DI students did not employ effective strategies during the consolidation phase so that they did not get more accurate prepositions. They did not have enough preparation for the subsequent instruction as they did not recognise the gaps. Despite some learning strategies, DI students did not show monitoring (metacognitive) strategies in learning.

Our conclusion on PF students' self-regulated learning is that a global learning goal was set right after the students were familiar with the instruction. Then, the learning sub-goals were constructed when doing the first communicative task, and the sub-goals were set more accurately by comparing and contrasting the first task performance to the description model in the consolidation provided the knowledge gaps were acknowledged. These more accurate individual learning goals led the students to use more effective learning strategies and metacognitive monitoring strategies when dealing with the instructional video. As a result, PF students outperformed DI students in the syntactic accuracy (i.e., procedural knowledge) of a written description of places and executed more self-regulated learning strategies and motivation. However, this conclusion needs further investigation, and further recommendations is covered in the next chapter.

CHAPTER 6 CONCLUSION AND IMPLICATION

6.1. Conclusions

This experimental research applied the concepts of productive failure in comparison with direct instruction in a university level of MUVE-based English language learning. The PF group finished a communicative task prior to a video-based instruction while the DI group watched a video-based instruction before finishing a communicative task. Another similar communicative task was assigned to both groups. A post-test on the written and spoken-based description of places revealed that there were differential results between PF and DI groups. The qualitative analysis on the students' behaviours while both groups accessed the video indicated a different result as well. The three findings suggest that 1) PF group outperformed DI group in both the overall score and the procedural knowledge (syntax) but performed equally in the declarative knowledge (vocabulary) in the written description of places; 2) PF group performed equally in all measures of oral description of places – overall score, declarative knowledge, and procedural knowledge; and 3) PF group enacted more metacognitive strategies compared to DI group during the video-based instruction.

The findings of this study indicate that the use of Productive Failure (PF) in EFL learning was effective to learn a foreign language – such as vocabulary and syntax in written-based assessment – carried out in production-based instruction. Looking it deeper, assigning a communicative task before instruction benefits students with their procedural knowledge learning in the written assessment.

PF learning mechanisms might work in EFL learning context in that the initial communicative task before instruction on vocabularies and syntax used in describing places enhances prior knowledge activation. When prior knowledge is activated, knowledge gaps will likely be acknowledged by comparing and contrasting students' performance and the canonical model in the consolidation phase. When this gap is filled during the instruction, the new linguistic schema will likely have resulted. Exposing to the similar but slightly more complicated task after the instruction will likely make a more solid schema. This process of learning is helpful for written-based syntax learning.

This study has proposed an alternative to MUVE-based EFL learning's pedagogical strategy to the common practice of using the instruct-practice approach similar to DI. While previous studies in MUVEs have reported their affordances in language learning (e.g., Huang, Grant, & Henderson, 2012; Mohammadi, 2017; Wang, 2015), this study contributes to using effective pedagogical strategy in MUVE-based EFL learning.

This study demonstrated that finishing an authentic task prior to instruction supported students' use of more meta-cognitive monitoring and cognitive learning strategies in the self-regulated learning framework. Prior knowledge activation and differentiation promoted in the first task helps sub-goal setting easier, which will likely guide cognitive learning strategies during the learning session. In personal-based online learning, self-regulated learning skill is an essential determinant of learning success.

Applying PF pedagogical strategy has a potential to improve lexical and syntactic accuracy in oral production. The reason was that though PF group performed equally in the oral assessment, they showed positive learning behaviour as indicated in their higher metacognitive strategies. From an oral production perspective, enforcing accuracy of oral output requires a considerable number of practices, at least two repeat tasks for fluency purposes (Bygate, 1999; Finardi, 2008) and at least four similar tasks repeated in a more extended period for accuracy purposes (Lambert et al., 2017). This study has given a meaningful contribution to enhancing accuracy and fluency in speaking, provided that both PF and DI produced comparable results with the same practice opportunity. It indicates that activity sequence may contribute to the development of lexical and syntactic accuracy in the target language's oral production apart from the number of practices.

6.2. Implications for English as a Foreign Language Learning

EFL practices have been mostly classroom-based, in which a teacher presents the material and students practice the target expressions with their classmates. Task-based language teaching has been the most widely used approach (Hu, 2005; Kiernan, 2005; Todd, 2006; Vilches, 2003) in Asian countries and the ideal tool in language teaching (Ellis, 2009; Robinson, 2011). In the implementation, however, the weak form application of TBLT has been the most common in the EFL context. Despite the reported success to increase fluency, accuracy, and complexity (e.g., Ho, 2017), authenticity has been a challenge in the EFL classroom (see Ozverir et al., 2016). Fortunately, studies have proven the effectiveness of a multi-user virtual environment (MUVE) as an alternative to overcome the authenticity issue (Jauregi et al., 2011). This research is a complement of the MUVE-based EFL practices in the pedagogical strategies.

The main findings of this study—the use of PF was more effective for students to learn written-based syntax (procedural) and supported more self-regulated learning strategies than the use of DI—contribute to English as a foreign language learning. The research results can be applied to design the activities in both classroom-based and online learning platforms. An authentic task should be created prior to the subsequent instruction to help students

activate and differentiate their prior knowledge. If prior knowledge has been activated, students can focus on which part of the instruction to pay attention to as they have acknowledged their gaps. This way, we can give the correct time and amount of guidance that they need.

6.2.1. Classroom-based EFL

A single production is widely adopted in the classroom-based EFL, in which EFL learners produce the target language in an open context after instruction and practice. Despite the reported effectiveness of improving communication skills, students' expressions during the production phase often rely too much on the model presented in the instruction (Ellis, 2018; Harris, 2018). There is also a possibility that students ignore the teacher's instruction due to lack of need or familiarity with the materials. The application of PF pedagogical strategy can be an alternative to the long history of task-based language learning application in the classroom that puts production (authentic task) after the instruction and guided practice (Littlewood, 2007; Vilches, 2003), especially in the weak-form of TBLT application.

Students can be given a communicative task at the onset of the instruction to activate the expressions needed to finish the task. They will likely fail to complete the task, and that is what we want. We let them fail so that they understand how far their prior knowledge can tackle the task. This way, they can acknowledge their knowledge gaps to be filled during the instruction. The instruction can be presented in the form of instructional video or teacher's presentation as long as students can compare and contrast their performance in the task and the expected performance from a model. The explicit instruction enables students to fill in their knowledge gaps to build substantial new knowledge of their own. It is expected that students will understand more profound about the target expressions expected from the communicative task.

6.2.2. Online based EFL

The authentic task prior to instruction in online learning can help the students focus on the instruction, either video-based or conference-based. It is widely noted that online learning has a high attrition rate (Gütl, Rizzardini, Chang, & Morales, 2014), partly due to students' motivation or self-regulated learning skills. Lack of motivation can be due to the students' inability to see the instruction's value as they do not know precisely what they need, i.e., knowledge gaps.

The finding that PF students enacted more effective self-regulated learning strategies is also an essential contribution to an online learning platform. A careful correlational study needs carrying out to explore the effect of activity sequencing on SRL development. Previous

studies have revealed that scaffolding could instil self-regulation among online learners (e.g., Azevedo et al., 2017; Taub et al., 2019; Taub, Mudrick, & Azevedo, 2018). Online learners do need SRL skills as they are physically distant from the teacher/instructor. PF can help students set their sub-goals once they have activated their prior knowledge from the first communicative task. Their sub-goal can be more accurate once they can monitor their learning by comparing their performance in the first task to the canonical one in the consolidation. Having more accurate sub-goals, students can determine precise learning strategies during instruction, which will help them fill in the knowledge gaps. Once the gaps are filled, students can have a stable schema, meaning that they have learned new things.

These findings are crucial as currently there is extensive investigation on MUVE-based EFL learning based on principles similar to DI pedagogical strategies (e.g., Rahayu & Jacobson, 2012; Wang, Calandra, Hibbard, & McDowell Lefaiver, 2012; Wang, Petrina, & Feng, 2017). This research is an innovation on the alternative pedagogical strategies, in which additional authentic communication before the instruction can be assigned in the commonly used task-based language learning.

6.3. Future Research Recommendations

This study is expected to contribute to MUVE-based EFL learning by applying PF pedagogical strategy—giving communicative tasks prior to instruction. Findings suggested that both PF and DI students gained equal declarative knowledge in written and spoken language production. PF students outperformed DI students in the procedural knowledge of the written output of the target knowledge but performed equally on the procedural knowledge of spoken production of the target language. Besides, PF students enacted more meta-cognitive monitoring and cognitive learning strategies during the instruction session than DI students. Despite those positive findings, limitations are presented in the next paragraphs and then followed by future research recommendations.

This is the first research on PF applied in language learning, to the best of my knowledge. Despite the positive results on procedural knowledge and self-regulated learning, replication is needed to ensure that the findings were not accidental. Replications of this research—either with the same MUVE or in classroom-basis—with the same pedagogical design would be necessary before it is applied in the classroom. If the findings are consistent with this research, it will strengthen the contribution of the PF pedagogical strategy to EFL learning. Bigger sample size is also suggested for more robust results.

Similar studies—either with the same MUVE or in a classroom-basis—could be carried out with modifications on the focus (e.g., conceptual knowledge or transfer). These

focuses will broaden the scope of PF effects on EFL learning and be in line with PF studies carried out in science and mathematics. Another modification might be in the duration. While the most effective (i.e., the largest effect size) experimental studies in technology-mediated EFL program should take 3-4 weeks (Chiu, 2013b; Tsai & Tsai, 2018), this research took only 80-90 minutes per individual in single participation. It will be interesting to see PF's effect in a more extended period of EFL learning (more practices).

PF and DI students equally improved students' vocabulary and syntax used in oral production. Previous studies on PF suggested that the number of practices that influenced procedural fluency, not the sequence of activities (Loibl & Rummel, 2014a; 2014b). In a language learning context, four practices after instruction are suggested to automatise the newly-learned vocabularies or expressions for accuracy (Lambert et al., 2017) and two practices for fluency (Bygate, 1999; Finardi, 2008). Therefore, modifying PF pedagogical design with more practices after the consolidation is advisable to explore its efficacy for foreign language learning.

The self-regulated learning results are limited to the learning strategies and monitoring strategies due to the focus on observable behaviour during the instruction. More comprehensive and correlational studies are needed to confirm and broaden the results of recent research. More complete SRL phases should be investigated, i.e., from the planning phase to the adaptation phase so that the effect of PF pedagogical strategies on SRL strategies will be more precise. These recommended future studies can be conducted in EFL or other contexts, given that this research was one of the first studies to investigate SRL strategies between PF and DI students.

Bibliography

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, *21*(3), 199–226. https://doi.org/10.1080/09588220802090246

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*. https://doi.org/10.1177/1362168810383329

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does Discovery-Based Instruction Enhance Learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

Ali Mohsen, M. (2016). The Use of Computer-Based Simulation to Aid Comprehension and Incidental Vocabulary Learning. *Journal of Educational Computing Research*. https://doi.org/10.1177/0735633116639954

Allaw, E., & McDonough, K. (2019). The effect of task sequencing on second language written lexical complexity, accuracy, and fluency. *System*, *85*, 102–104. https://doi.org/10.1016/j.system.2019.06.008

Anderson, R. C., Hiebert, E. H., Scott, J. A., Wilkinson, I. A. G., Anderson, R. C., Hieb, E. H., … Becker, W. (1988). Becoming a nation of readers: The reaport of the commission on reading. *Education and Treatment of Children*, *11*(4), 389–396. Retrieved from http://jstor.com/stable/42899086

Ashman, G., Kalyuga, S., & Sweller, J. (2020). Problem-solving or Explicit Instruction: Which Should Go First When Element Interactivity Is High? *Educational Psychology Review*, *32*, 229–247. https://doi.org/10.1007/s10648-019-09500-5

Azevedo, R. (2018). Using hypermedia as a metacognitive tool for enhancing student learning? the role of self-regulated learning. *Computers as Metacognitive Tools for Enhancing Learning: A Special Issue of Educational Psychologist*, *40*(4), 199–209. https://doi.org/10.4324/9781315866239-2

Azevedo, R., Behnagh, R. F., Harley, J. M., & Trevors, G. (2010). Learning in Student-Centered Learning Environments. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (pp. 171–197). NY: Routledge.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated Learning with MetaTutor: Advancing the Science of Learning with MetaCognitive Tools. In M. S. Khine & I. M. Saleh (Eds.), *New Science of Learning: Cognition, Computers and*

Collaboration in Education (pp. 225–247). https://doi.org/10.1007/978-1-4419-5716-0

Azevedo, R., Johnson, A., Chauncey, A., & Graesser, A. (2015). Use of Hypermedia to Assess and Convey Self-Regulated Learning. In *Handbook of Self-Regulation of Learning and Performance*. https://doi.org/10.4324/9780203839010.ch7

Azevedo, R., Landis, R. S., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J. M., … Hossain, G. (2012). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7315 LNCS*, 212–221. https://doi.org/10.1007/978-3-642-30950-2_27

Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist*, *45*(4), 210–223. https://doi.org/10.1080/00461520.2010.515934

Azevedo, R., Taub, M., Mudrick, N. V., Millar, G. C., Bradbury, A. E., & Price, M. J. (2017). Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies. *Informational Environments: Effects of Use, Effective Designs*, 225–247. https://doi.org/10.1007/978-3-319-64274-1_10

Bachman, L., & Palmer, A. (2010). *Language Assessment: Developing Language Assessment and Justifying their Use in the Real World*. New York: Oxford University Press.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: WH Freeman.

Baralt, M., Gilabert, R., & Robinson, P. (2014). *Task Sequencing and Instructed Second Language Learning*. https://doi.org/10.5040/9781472593726.ch-001

Baralt, M., & Gomez, J. M. (2017). Task-based language teaching online: A guide for teachers. *Language Learning & Technology*, *21*(213), 28–43. Retrieved from http://llt.msu.edu/issues/october2017/baraltmorcillogomez.pdf

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

Barcroft, J. (2009). Effects of Synonym Generation on Incidental and Intentional L2 Vocabulary Learning During Reading. *TESOL Quarterly*, *43*(1), 79–103. https://doi.org/10.1002/j.1545-7249.2009.tb00228.x

Baumann, J. F. (1988). Direct Instruction Reconsidered. *Journal of Reading*, *31*(8), 712–718. Retrieved from http://www.jstor.com/stable/40032953

Bock, J. K. (1986). Meaning, Sound, and Syntax. Lexical Priming in Sentence Production.

Journal of Experimental Psychology: Learning, Memory, and Cognition. https://doi.org/10.1037/0278-7393.12.4.575

Bock, J. K., & Levelt, W. (1994). Language production: grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego: Academic Press.

Borreguero Zuloaga, M., & De Marco, A. (2020). The Role of Immersion and Non-immersion Contexts in L2 Acquisition: A Study Based on the Analysis of Interactional Discourse Markers. In *Corpus Pragmatics*. https://doi.org/10.1007/s41701-020-00093-x

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking Transfer: a Simple Proposal with Multiple Implications. *Review of Research in Education*, *24*, 61–100. Retrieved from https://www.jstor.org/stable/1167267

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: Comparing task-specific self-reports to traces. *Metacognition and Learning*, *2*(1), 1–20. https://doi.org/10.1007/s11409-007-9004-y

Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*. https://doi.org/10.1177/136216889900300302

Bygate, M. (2013). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing* (pp. 23–48). London: New York: Routledge.

Cacciamani, S., Cesareni, D., Martini, F., Ferrini, T., & Fujita, N. (2012). Influence of participation, facilitator styles, and metacognitive reflection on knowledge building in online university courses. *Computers and Education*, *58*(3), 874–884. https://doi.org/10.1016/j.compedu.2011.10.019

Cahyono, B. Y. (2008). The Teaching of EFL Vocabulary in the Indonesian Context: The State Of The Art. *TEFLIN Journal*, *19*(1), 1–17. https://doi.org/10.15639/teflinjournal.v19i1/1-17

Cairns, P., Cox, A., & Nordin, A. I. (2014). Immersion in Digital Games: Review of Gaming Experience Research. In M. C. Angelides & H. Agius (Eds.), *Handbook of Digital Games* (pp. 339–361). Retrieved from http://www.vwh-verlag.de/vwh/wp-content/uploads/2007/10/pacher_inhaltsverzeichnis.pdf

Carey, S. (1999). Sources of conceptual change. In E. K. Scholnick, K. Nelson, & P. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 293–326). Mahwah, NJ: Lawrence Erlbaum Associates.

Carless, D. (2002). Implementing Task-based Learning with Young Learners. *ELT Journal*, *56*(4), 389–396. Retrieved from http://eltj.oxfordjournals.org/

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. Retrieved from https://www.jstor.org/stable/3233596

Chen, D. (2010). Enhancing the learning of Chinese with Second Life. *Journal of Technology and Chinese Language Teaching*. Retrieved from http://www.tclt.us/journal/2010v1n1/chen.pdf

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

Cheng, L., Li, M., Kirby, J. R., Qiang, H., & Wade-Woolley, L. (2010). English language immersion and students' academic achievement in English, Chinese and mathematics. *Evaluation and Research in Education*. https://doi.org/10.1080/09500790.2010.489150

Chi, M. T. H. (2000). Self-Explaining Expository Texts: The Dual Processes of Generating Inferences and Repairing Mental Models. In R. Glaser (Ed.), *Advances in Instructional Psychology* (pp. 161–238). Retrieved from https://doi-org.ezproxy.library.sydney.edu.au/10.4324/9780203774526

Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, *32*(2), 301–341. https://doi.org/10.1080/03640210701863396

Chiu, Y. H. (2013a). Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology*, *44*(2), 52–57. https://doi.org/10.1111/j.1467-8535.2012.01342.x

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

Chiu, Y. H., Kao, C. W., & Reynolds, B. L. (2012). The relative effectiveness of digital game-based learning types in English as a foreign language setting: A meta-analysis. *British Journal of Educational Technology*, *43*(4), 104–107. https://doi.org/10.1111/j.1467-8535.2012.01295.x

Chowrira, S. G., Smith, K. M., Dubois, P. J., & Roll, I. (2019). DIY productive failure: boosting performance in a large undergraduate biology course. *Npj Science of Learning*. https://doi.org/10.1038/s41539-019-0040-6

Cohen, J. (1977). F Tests on Means in the Analysis of Variance and Covariance. *Statistical Power Analysis for the Behavioral Sciences*, 273–406. https://doi.org/10.1016/b978-0-12-179060-8.50013-x

Cook, V. (2008). *Second Language Learning and Teaching* (4th ed). London: Hodder Education.

Cooke-Plagwitz, J. (2008). New directions in CALL: An objective introduction to second life. *CALICO Journal*. https://doi.org/10.1558/cj.v25i3.547-557

Council of Europe. (2001). the Common European Framework of Reference for Languages : Learning, Teaching, Assessment. *Council of Europe*. https://doi.org/10.1017/S0267190514000221

Cowan, N., & AuBuchon, A. M. (2008). Short-term memory loss over time without retroactive stimulus interference. *Psychonomic Bulletin and Review*, *15*(1), 230–235. https://doi.org/10.3758/PBR.15.1.230

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*(2), 428–434. https://doi.org/10.1037/0022-0663.94.2.428

Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. In *Educational Research*. https://doi.org/10.1017/CBO9781107415324.004

Creswell, J. W., & Creswel, J. D. (2018). *Research Design: Qualitative, quantitative, and mixed methods approaches* (Fifth). Thousand Oaks, CA: SAGE Publications, Inc.

Criado, R. (2013). A critical review of the Presentation-Practice-Production Model (PPP) in Foreign Language Teaching. In R. Monroy (Ed.), *Homenaje a Francisco Gutiérrez Diez* (pp. 97–115). Murcia: Edit.um.

De Lucia, A., Francese, R., Passero, I., & Tortora, G. (2009). Development and evaluation of a virtual campus on Second Life: The case of SecondDMI. *Computers and Education*, *52*(1), 220–233. https://doi.org/10.1016/j.compedu.2008.08.001

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology*, *113*(4), 552–568. https://doi.org/10.1016/j.jecp.2012.06.009

DeHaan, J., Michael Reed, W. M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning and Technology*, *14*(2), 74–94. Retrieved from

http://llt.msu.edu/vol14num2/dehaanreedkuwada.pdf%0AJune

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542. https://doi.org/10.1207/s15516709cog2304_6

Dickey, M. D. (2005). Three-dimensional virtual worlds and distance learning: two case studies of Active Worlds as a medium for distance education. *British Journal of Educational Psychology*, *36*(3), 439–451. https://doi.org/10.1016/B978-0-12-813467-2.00063-8

Dickey, M. D. (2011). The pragmatics of virtual worlds for K-12 educators: Investigating the affordances and constraints of Active Worlds and Second Life with K-12 in-service teachers. *Educational Technology Research and Development*, *59*(1), 1–20. https://doi.org/10.1007/s11423-010-9163-4

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

diSessa, A. A. (2014). A history of conceptual change research: Threads and fault lines. *The Cambridge Handbook of the Learning Sciences, Second Edition*, 88–108. https://doi.org/10.1017/CBO9781139519526.007

Doerr, H. M., & Lesh, R. (2003). A Modelling Perspective on Teacher Development. In R. Lesh & H. M. Doerr (Eds.), *Beyond Constructivism* (pp. 125–140). Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Z. (2005). Language Learning Strategies and Student Self-Regulation. In Z. Dornyei (Ed.), *The psychology of the language Learner: Individual differences in second language acquisition* (pp. 162–196). Retrieved from https://ebookcentral-proquest-com.ezproxy.library.sydney.edu.au/lib/usyd/detail.action?docID=257290.

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338–348. https://doi.org/10.1016/j.chb.2015.05.041

Dumford, A. D., & Miller, A. L. (2018). Online learning in higher education: exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education*, *30*(3), 452–465. https://doi.org/10.1007/s12528-018-9179-z

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, *19*(3), 221–246.

https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ellis, R. (2018). Reflections on task-based language teaching. In *Reflections on Task-Based Language Teaching*. https://doi.org/10.21832/ELLIS0131

Ellis, R. (2019). Towards a modular language curriculum for using tasks. *Language Teaching Research*, *23*(4), 454–475. https://doi.org/10.1177/1362168818765315

Engelmann, S. (2007). *Teaching Needy Kids in Our Backward System: 42 Years of Trying*. Eugene, Or: ADI Press.

Ergen, B., & Kanadli, S. (2017). The Effect of Self-regulated Learning Strategies on Academic Achievement: A Meta-analysis Study. *Eurasian Journal of Educational Research*, *69*, 55–74. Retrieved from http://dx.doi.org/10.14689/ejer.2017.69.4

Finardi, K. R. (2008). Effects of task repetition on L2 oral performance. *Trabalhos Em Linguística Aplicada*. https://doi.org/10.1590/s0103-18132008000100003

Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, *20*(3), 321–340. https://doi.org/10.1177/1362168815570142

Gaille, L. (2017). 16 Advantages and Disadvantages of Experimental Research. Retrieved December 24, 2020, from Ittana.org website: https://vittana.org/16-advantages-and-disadvantages-of-experimental-research

Gavriel, J. (2015). Cognitive apprenticeship. *Education for Primary Care*, *26*(6), 422–423. https://doi.org/10.1080/14739879.2015.1101851

Girvan, C. (2018). What is a virtual world? Definition and classification. *Educational Technology Research and Development*, *66*(5), 1087–1100. https://doi.org/10.1007/s11423-018-9577-y

Goda, Y., Yamada, M., Kato, H., Matsuda, T., Saito, Y., & Miyagawa, H. (2015). Procrastination and other learning behavioral types in e-learning and their relationship with learning outcomes. *Learning and Individual Differences*, *37*, 72–80. https://doi.org/10.1016/j.lindif.2014.11.001

Godwin-Jones, R. (2011). Emerging Technologies: Autonomous language learning. *Language Learning and Technology*, *15*(3), 4–10. Retrieved from http://llt.msu.edu/issues/october2011/emerging.pdf

Graesser, A. C., Jeon, M., & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes*, *45*(4–5), 298–322. https://doi.org/10.1080/01638530802145395

Graesser, A. C., Mello, S. D., & Person, N. (2009). Meta-Knowledge in Tutoring. In D. J.

Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 361–382). New York: Routledge.

Graesser, Art, Penumatsa, P., Ventura, M., Cai, Z., & Hu, X. (2007). Using LSA in AutoTutor: Learning Through Mixed-Initiative Dialogue in Natural Language. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of Semantic Analysis* (pp. 243–262). New Jersey: Routledge.

Graesser, Arthur, & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*, *45*(4), 234–244. https://doi.org/10.1080/00461520.2010.515933

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research*, *77*(3), 334–372. https://doi.org/10.3102/003465430303953

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, *34*(1), 18–29. https://doi.org/10.1016/j.cedpsych.2008.05.006

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. *L@S 2014 - Proceedings of the 1st ACM Conference on Learning at Scale*. https://doi.org/10.1145/2556325.2566239

Gütl, C., Rizzardini, R. H., Chang, V., & Morales, M. (2014). Attrition in MOOC: Lessons Learned from Drop-Out Students. *Communications in Computer and Information Science*. https://doi.org/10.1007/978-3-319-10671-7_4

Hardiman, P. T., Pollatsek, A., & Well, A. D. (1986). Learning to Understand the Balance Beam. *Cognition and Instruction*, *3*(1), 63–86. Retrieved from http://www.jstor.com/stable/3233605

Harris, J. (2018). Responding to the critics: Implementation of TBLT in Japan. *Indonesian Journal of Applied Linguistics*, *8*(1), 139–148. https://doi.org/10.17509/ijal.v8i1.11473

Hmelo-silver, C. E. (2004). Problem-Based Learning : What and How Do Students Learn ? *Educational Psychology Review*, *16*(3), 235–266. Retrieved from https://doi.org/10.1080/00461520701263368

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Ho, R. C. C. (2017). Comparison of task repetition and task sequencing in terms of their effects on L2 students' oral performance. *International Journal of Instruction*, *10*(1), 307–322. https://doi.org/10.12973/iji.2017.10119a

Holmes, N. G., Day, J., Park, A. H. K., Bonn, D. A., & Roll, I. (2014). Making the failure more productive: scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. *Instructional Science*, *42*(4), 523–538. https://doi.org/10.1007/s11251-013-9300-7

Hong, S. W., Jeong, Y., Kalay, Y. E., Jung, S., & Lee, J. (2016). Enablers and barriers of the multi-user virtual environment for exploratory creativity in architectural design collaboration. *CoDesign*, *12*(3), 151–170. https://doi.org/10.1080/15710882.2015.1081239

Hu, G. (2005). English language education in China: Policies, progress, and problems. *Language Policy*, *4*(1), 5–24. https://doi.org/10.1007/s10993-004-6561-7

Huang, H., Grant, S., & Henderson, L. (2012). The impact of Chinese language lessons in a virtual world on university students' self-efficacy beliefs. *Australasian Journal of Educational Technology*, *28*(3), 400–419. https://doi.org/10.14742/ajet.842

Jacobson, M. J., Goldwater, M., Markauskaite, L., Lai, P. K., Kapur, M., Roberts, G., & Hilton, C. (2020). Schema abstraction with productive failure and analogical comparison: Learning designs for far across domain transfer. *Learning and Instruction*, *65*(March 2018), 101222. https://doi.org/10.1016/j.learninstruc.2019.101222

Jacobson, M. J., Kim, B., Pathak, S., & Zhang, B. H. (2015). To guide or not to guide: issues in the sequencing of pedagogical structure in computational model-based learning. *Interactive Learning Environments*, *23*(6), 715–730. https://doi.org/10.1080/10494820.2013.792845

Jacobson, M. J., Markauskaite, L., Portolese, A., Kapur, M., Lai, P. K., & Roberts, G. (2017). Designs for learning about climate change as a complex system. *Learning and Instruction*, *52*, 1–14. https://doi.org/10.1016/j.learninstruc.2017.03.007

Jacobson, M. J., Taylor, C. E., & Richards, D. (2016). Computational scientific inquiry with virtual worlds and agent-based models: new ways of doing science to learn science. *Interactive Learning Environments*, *24*(8), 2080–2108. https://doi.org/10.1080/10494820.2015.1079723

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 649–667. https://doi.org/10.1016/S0022-5371(78)90393-6

Jauregi, K., Canto, S., de Graaff, R., Koenraad, T., & Moonen, M. (2011). Verbal interaction in Second Life: Towards a pedagogic framework for task design. *Computer Assisted Language Learning*, *24*(1), 77–101. https://doi.org/10.1080/09588221.2010.538699

Jean, G., & Simard, D. (2013). Deductive versus inductive grammar instruction: Investigating possible relationships between gains, preferences and learning styles. *System*, *41*, 1023–1042. https://doi.org/10.1016/j.system.2013.10.008

Jupp, V. (2006). *The SAGE Dictionary of Social Research Methods* (Vols 1-10). https://doi.org/10.4135/9780857020116

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*(3), 379–424. https://doi.org/10.1080/07370000802212669

Kapur, M. (2011). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science*, *39*, 561–579. https://doi.org/10.1007/s11251-010-9144-3

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*, 651–672. https://doi.org/10.1007/s11251-012-9209-6

Kapur, M. (2014a). Comparing Learning From Productive Failure and Vicarious Failure. *Journal of the Learning Sciences*, *23*(4), 651–677. https://doi.org/10.1080/10508406.2013.819000

Kapur, M. (2014b). Productive failure in learning math. *Cognitive Science*, *38*(5), 1008–1022. https://doi.org/10.1111/cogs.12107

Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, *1*(1), 51–65. https://doi.org/10.1080/23735082.2015.1002195

Kapur, M. (2016). Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning. *Educational Psychologist*, *51*(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457

Kapur, M., & Bielaczyc, K. (2011). Classroom-based experiments in productive failure. *Proceedings of the 33rd Annual Conference of …*, 2812–2817. Retrieved from http://jmfs1.ortn.edu/MySchool/dhundermark/Links/Productive Failure Research - Kapor.pdf

Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *Journal of the Learning Sciences*, *21*(1), 45–83. https://doi.org/10.1080/10508406.2011.591717

Kapur, M., & Kinzer, C. K. (2009). Productive failure in CSCL groups. *International Journal*

of Computer-Supported Collaborative Learning, *4*(1), 21–46. https://doi.org/10.1007/s11412-008-9059-z

Kennedy-Clark, S., Jacobson, M., & Reimann, P. (2009). Productive failure in inquiry learning in a multi-user virtual environment. *ASCILITE 2009 - The Australasian Society for Computers in Learning in Tertiary Education*.

Khezrlou, S., & Ellis, R. (2017). Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *System*, *65*, 104–116. https://doi.org/10.1016/j.system.2017.01.009

Kiernan, P. (2005). Storytelling with Low-levelLearners: Developing NarrativeTasks. In E. Corony & J. Willis (Eds.), *Teachers ExploringTasks in English Language Teaching* (pp. 58–68). Macmillan: Palgrave MacMillan.

Kim, N. (2020). The effects of different task sequences on novice L2 learners' oral performance in the classroom. *Language Teaching Research*. https://doi.org/10.1177/1362168820937548

Kim, Y. J., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, *41*, 829–840. https://doi.org/10.1016/j.system.2013.08.005

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, *41*(March 2015), 87–98. https://doi.org/10.1207/s15326985ep4102

Kirschner, P., & Clark, R. E. (2004). Why Discovery Learning, Problem-Based Learning, Experiential Learning and Inquiry-Based Learning Do Not Work. *Educational Psychologist*.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Knell, E., Haiyan, Q., Miao, P., Yanping, C., Siegel, L. S., Lin, Z., & Wei, Z. (2007). Early English immersion and literacy in Xi'an, China. *Modern Language Journal*, *91*(iii), 395–417. https://doi.org/10.1111/j.1540-4781.2007.00586.x

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*, 239–264. https://doi.org/10.1007/s10648-007-9049-0

Kolb, D. (1984). *Experiential learning : experience as the source of learning and*

development. Prentice-Hall.

Kolb, D. (2014). *Experiential learning: experience as the source of learning and development* (Second). Pearson FT Press.

Kontogeorgiou, A. M., Bellou, J., & Mikropoulos, T. A. (2008). Being inside the Quantum Atom. *PsychNology Journal*, *6*(1), 83–98. Retrieved from http://ezproxy.library.usyd.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=31971220&site=ehost-live

Kumaravadivelu, B. (2006). *Understanding language teaching: From method to postmethod*. https://doi.org/10.4324/9781410615725

Lai, P. K. (2017). Learning Nanoscience from a nanoparticle's perspective: A computationally embodies learning experience. The University of Sydney.

Lai, P. K., Jacobson, M. J., & Goldwater, M. (2018). Learning nanoscience concepts through a nanoscale experience. *Proceedings of International Conference of the Learning Sciences, ICLS*.

Lai, P. K., Portolese, A., & Jacobson, M. J. (2017). Does sequence matter? Productive failure and designing online authentic learning for process engineering. *British Journal of Educational Technology*, *48*(6), 1217–1227. https://doi.org/10.1111/bjet.12492

Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, *39*, 167–196. https://doi.org/10.1017/S0272263116000085

Lambert, C., & Robinson, P. (2014). Learning to Perform Narrative Tasks: A Semester-long Classroom Study of L2 Task Sequencing Effects. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task Sequencing and Instructed Second Language Learning* (pp. 207–230). https://doi.org/10.5040/9781472593726.ch-008

Lan, Y. J. (2014). Does second life improve mandarin learning by overseas chinese students? *Language Learning and Technology*, *18*(2), 36–56. Retrieved from http://llt.msu.edu/issues/june2014/action2.pdf%0AJune

Lan, Y. J., Kan, Y. H., Hsiao, I. Y. T., Yang, S. J. H., & Chang, K. E. (2013). Designing interaction tasks in second life for Chinese as a foreign language learners: A preliminary exploration. *Australasian Journal of Educational Technology*, *29*(2), 184–202. https://doi.org/10.14742/ajet.144

Lee, H. S., & Anderson, J. R. (2013). Student Learning: What Has Instruction Got to Do With It? *Annual Review of Psychology*, *64*(1), 445–469. https://doi.org/10.1146/annurev-psych-113011-143833

Levelt, W. J.M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13464–13471. https://doi.org/10.1073/pnas.231459498

Levelt, Willem J.M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/S1364-6613(99)01319-4

Li, S., Ellis, R., & Zhu, Y. (2016). Task-Based Versus Task-Supported Language Instruction: An Experimental Study. *Annual Review of Applied Linguistics*, *36*, 205–229. https://doi.org/10.1017/S0267190515000069

Lin, S. Y., Aiken, J. M., Seaton, D. T., Douglas, S. S., Greco, E. F., Thoms, B. D., & Schatz, M. F. (2017). Exploring physics students' engagement with online instructional videos in an introductory mechanics course. *Physical Review Physics Education Research*, *13*, 020138-1-020138–18. https://doi.org/10.1103/PhysRevPhysEducRes.13.020138

Liou, H. C. (2012). The roles of Second Life in a college computer-assisted language learning (CALL) course in Taiwan, ROC. *Computer Assisted Language Learning*, *25*(4), 365–382. https://doi.org/10.1080/09588221.2011.597766

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, *40*(3), 243–249. https://doi.org/10.1017/S0261444807004363

Littlewood, W. (2013). Developing a Context-sensitive Pedagogy for Communication-oriented Language Teaching. *English Teaching*, *68*(3), 3–25. https://doi.org/10.15858/engtea.68.3.201309.3

Loehr, A. M., Fyfe, E. R., & Rittle-Johnson, B. (2014). Wait for it . . . delaying instruction improves mathematics problem solving: A classroom study. *Journal of Problem Solving*, *7*, 36–49. https://doi.org/10.7771/1932-6246.1166

Loibl, K., & Leuders, T. (2019). How to make failure productive: Fostering learning from errors through elaboration prompts. *Learning and Instruction*. https://doi.org/10.1016/j.learninstruc.2019.03.002

Loibl, K., & Rummel, N. (2014a). Knowing what you don't know makes failure productive. *Learning and Instruction*, *34*, 74–85. https://doi.org/10.1016/j.learninstruc.2014.08.004

Loibl, K., & Rummel, N. (2014b). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science*, *42*(3), 305–326. https://doi.org/10.1007/s11251-013-9282-5

Loibl, K., & Rummel, N. (2015). Productive Failure as strategy against the double curse of incompetence. *Learning: Research and Practice*, *1*(2), 113–121.

https://doi.org/10.1080/23735082.2015.1071231

Malicka, A. (2014). The Role of Task Sequencing in Monologic Oral Production. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task Sequencing and Instructed Second Language Learning* (pp. 71–94). London: Bloomsbury Academic.

Mar-molinero, C., & Stevenson, P. (2006). Language ideologies, policies and practices: language and the future of europe. In *Journal of Chemical Information and Modeling*. https://doi.org/10.1017/CBO9781107415324.004

Marklund, B. B., Backlund, P., Dahlin, C. J., Engström, H., & Wilhelmsson, U. (2014). A Game-Based Approach to Support Social Presence and Awareness in Distributed Project-Based Learning. *International Journal of Game-Based Learning*, *4*(1), 1–20. https://doi.org/10.4018/IJGBL.2014010101

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice*, *41*(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule against Pure Discovery Learning? The Case for Guided Methods of Instruction. *American Psychologist*, *59*(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14

McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning*, *10*(1), 43–75. https://doi.org/10.1007/s11409-014-9132-0

McLaren, B. M., Van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behavior*, *55*, 87–99. https://doi.org/10.1016/j.chb.2015.08.038

Mikropoulos, T. A. (2006). Presence: A unique characteristic in educational virtual environments. *Virtual Reality*, *10*(3–4), 197–206. https://doi.org/10.1007/s10055-006-0039-1

Mikum, S., Suksakulchai, S., Chaisanit, S., & Murphy, E. (2018). Students' participation in peer-to-peer communication supported by social media. *Education and Information Technologies*, *23*(2), 659–679. https://doi.org/10.1007/s10639-017-9628-8

Mohammadi, F. S. (2017). The Effect of Authentic Problem-Based Vocabulary Tasks on Vocabulary Learning of EFL Learners. *International Journal of Education and Literacy Studies*, *5*(3), 35. https://doi.org/10.7575//aiac.ijels.v.5n.3p.35

Moos, D. C., & Azevedo, R. (2008a). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology*, *33*, 270–298.

https://doi.org/10.1016/j.cedpsych.2007.03.001

Moos, D. C., & Azevedo, R. (2008b). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology*. https://doi.org/10.1016/j.cedpsych.2007.03.001

Myers, J. L., & Chang, S.-F. (2009). A Multiple-Strategy-Based Approach to Word and Collocation Acquisition. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *47*(2), 179–207. Retrieved from https://doi.org/10.1515/iral.2009.008

Nelson, B., Ketelhut, D. J., Clarke, J., Bowman, C., & Dede, C. (2005). Design-Based Research Strategies for Developing a Scientific Inquiry Curriculum in a Multi-User Virtual Environment. *Educational Technology*, *45*(1), 21–27. Retrieved from http://www.jstor.com/stable/44429185

Newman, P. M., & DeCaro, M. S. (2019). Learning by exploring: How much guidance is optimal? *Learning and Instruction*, *62*(November 2018), 49–63. https://doi.org/10.1016/j.learninstruc.2019.05.005

Noh, G. O., & Kim, D. H. (2019). Effectiveness of a self-directed learning program using blended coaching among nursing students in clinical practice: A quasi-experimental research design. *BMC Medical Education*, *19*(1), 1–8. https://doi.org/10.1186/s12909-019-1672-1

Ozverir, I., Herrington, J., & Osam, U. V. (2016). Design principles for authentic learning of English as a foreign language. *British Journal of Educational Technology*, *47*(3), 484–493. https://doi.org/10.1111/bjet.12449

Paechter, M., & Maier, B. (2010). Online or face-to-face? Students' experiences and preferences in e-learning. *Internet and Higher Education*, *13*(4), 292–297. https://doi.org/10.1016/j.iheduc.2010.09.004

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages*. Retrieved from https://ebookcentral-proquest-com.ezproxy1.library.usyd.edu.au/lib/usyd/detail.action?docID=622423

Pathak, S. A., Kim, B., Jacobson, M. J., & Zhang, B. (2011). Learning the physics of electricity: A qualitative analysis of collaborative processes involved in productive failure. *International Journal of Computer-Supported Collaborative Learning*, *6*, 57–73. https://doi.org/10.1007/s11412-010-9099-z

Paulus, C. T. M., Lester, J. N., & Dempster, P. G. (2014). Transcribing Audio and Video Data. In *Digital Tools for Qualitative Research* (pp. 93–114). https://doi.org/https://dx.doi.org/10.4135/9781473957671 Print

Pea, R. D. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *The Journal of the Learning Sciences*, *13*(3), 423–451. https://doi.org/10.1207/s15327809jls1303

Pellas, N. (2014). The influence of computer self-efficacy, metacognitive self-regulation and self-esteem on student engagement in online learning programs: Evidence from the virtual world of Second Life. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2014.02.048

Peterson, M. (2011). Towards a Research Agenda for the Use of Three-Dimensional Virtual Worlds in Language Learning. *CALICO Journal*, *29*(1), 67–80. Retrieved from https://doi.org/10.11139/cj.29.1.67-80

Phakiti, A. (2014). A Hybrid Approach for Experimental Research. In A. Phakiti (Ed.), *Experimental Research Methods in Language Learning* (pp. 143–157). London: Bloomsbury Academic.

Piaget, J. (2006). *The Origin of Intelligenge in the Child* (1953rd ed.). https://doi.org/10.1017/CBO9781107415324.004

Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In Monique Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 451–502). https://doi.org/10.1016/b978-012109890-2/50043-3

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning*, *53*(2), 233–284. https://doi.org/10.1111/1467-9922.00217

Rahayu, P., & Jacobson, M. J. (2012). Speaking self-efficacy and english as a foreign language: Learning processes in a multi-user virtual environment. In J. Peterson, O. Lee, T. Islam, & M. Piscioneri (Eds.), *Effectively Implementing Information Communication Technology in Higher Education in the Asia-Pacific Region* (pp. 161–182). New York: Nova Science Publishers.

Reynolds, B. L. (2015). The effects of word form variation and frequency on second language incidental vocabulary acquisition through reading. *Applied Linguistics Review*, *6*(4), 467–497. https://doi.org/10.1515/applirev-2015-0021

Rice, C. A., & Tokowicz, N. (2020). A review of laboratory studies of adult second language vocabulary training. *Studies in Second Language Acquisition*, *42*(2), 439–470.

https://doi.org/10.1017/S0272263119000500

Richards, J. C., & Reppen, R. (2014). Towards a pedagogy of grammar instruction. *RELC Journal*, *45*(1), 5–25. https://doi.org/10.1177/0033688214522622

Rittle-Johnson, B., & Schneider, M. (2015). Developing Conceptual and Procedural Knowledge of Mathematics. In R. Kadosh & A. Dowker (Eds.), *Oxford Handbook of Numerical Cognition* (pp. 1118–1134). Retrieved from https://doi.org/10.1093/oxfordhb/9780199642342.013.014

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing Conceptual and Procedural Understanding. *Jounral of Educational Psychology*, Vol. 91, pp. 346–362. https://doi.org/10.1037//0022-0663.93.2.346

Robinson, P. (2006). Criteria for classifying and Sequencing Pedagogic Tasks. In M. del P. G. Mayo (Ed.), *Investigating Tasks in Formal Language Learning* (pp. 7–26). https://doi.org/10.21832/9781853599286-004

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Pütz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind.* (pp. 243–268). Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc7&NEWS=N&AN=2010-14148-013

Robinson, P. (2011). Task-Based Language Learning: A Review of Issues. *Language Learning*, *61*(1), 1–36. https://doi.org/10.1111/j.1467-9922.2011.00641.x

Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and Directions in the Development of TBLT* (pp. 87–122). https://doi.org/10.1075/tblt.8.04rob

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL - International Review of Applied Linguistics in Language Teaching*, *45*(3), 161–176. https://doi.org/10.1515/iral.2007.007

Rovers, S. F. E., Clarebout, G., Savelberg, H. H. C. M., de Bruin, A. B. H., & van Merriënboer, J. J. G. (2019). Granularity matters: comparing different ways of measuring self-regulated learning. *Metacognition and Learning*, *14*(1), 1–19. https://doi.org/10.1007/s11409-019-09188-6

Royer, J. M. (1979). Theories of the transfer of learning. *Educational Psychologist*, *14*(1), 53–69. Retrieved from https://doi.org/10.1080/00461527909529207

Rugasken, K., & Harris, J. A. (2009). English Camp: A Language Immersion Program in Thailand. *The Learning Assistance Review*, *14*(2), 43–51. Retrieved from https://link.gale.com/apps/doc/A260332635/AONE?u=usyd&sid=AONE&xid=2d489eca

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *Modern Language Journal*, *90*(2), 228–243. https://doi.org/10.1111/j.1540-4781.2006.00394.x

Saito, K. (2014). The Role of Age of Acquisition in Late Second Language Oral Proficiency. *Studies in Second Language Acquisition*, *37*(4), 713–743. https://doi.org/10.1017/S0272263115000248

Salkind, N. J. (2010). Kr-20. In *Encyclopedia of research design* (pp. 668–668). https://doi.org/10.4135/9781412961288.n205

Savage, B. L., & Hughes, H. Z. (2014). How does Short-term Foreign Language Immersion Stimulate Language Learning? *Frontiers: The Interdisciplinary Journal of Study Abroad*, *XXIV*, 103–120. https://doi.org/10.36366/frontiers.v24i1.339

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–5223. https://doi.org/10.1207/s1532690xci1604_4

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129–184. https://doi.org/10.1207/s1532690xci2202_1

Segers, E., & Verhoeven, L. (2003). Effects of vocabulary training by computer in kindergarten. *Journal of Computer Assisted Learning*, *19*(April), 557–566. Retrieved from https://doi.org/10.1046/j.0266-4909.2003.00058.x

Shintani, N. (2015). The Effectiveness of Processing Instruction and Production-based Instruction on L2 Grammar Acquisition: A Meta-Analysis. *Applied Linguistics*, *36*(3), 306–325. https://doi.org/10.1093/applin/amu067

Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning*, *63*(2), 296–329. https://doi.org/10.1111/lang.12001

Sinha, T., & Kapur, M. (2019). When Productive Failure Fails. *41st Annual Conference of the Cognitive Science Society*, (May), 2811–2817.

Sleeman, D., Kelly, A. E., Martinak, R., Ward, R. D., & Moore, J. L. (1989). Studies of diagnosis and remediation with high school algebra students. *Cognitive Science*, *13*(4), 551–568. https://doi.org/10.1016/0364-0213(89)90023-2

Steinkuehler, C. A. (2006). Massively Multiplayer Online Video Gaming as Participation in a Discourse. *Mind, Culture, and Activity*, *13*(1), 38–52. https://doi.org/10.1207/s15327884mca1301_4

Sundqvist, P., & Wikström, P. (2015). Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System*, *51*, 65–76. https://doi.org/10.1016/j.system.2015.04.001

Sweller, J., Kirschner, P. A., & Clark, R. E. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist*, *42*(2), 115–121. https://doi.org/10.1080/00461520701263426

Tammenga-Helmantel, M., Arends, E., & Canrinus, E. T. (2014). The effectiveness of deductive, inductive, implicit and incidental grammatical instruction in second language classrooms. *System*, *45*, 198–210. https://doi.org/10.1016/j.system.2014.06.003

Tarone, E. (2005). Speaking in a Second Language. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 483–502). Mahwah, NJ: Lawrence Erlbaum.

Taub, M., & Azevedo, R. (2017). How does Prior Knowledge influence Fixations On and Sequences Of Cognitive and Metacognitive SRL Processes during Learning with an ITS? *International Journal of Artificial Intelligence in Education*, *29*(1), 1–28. Retrieved from https://doi.org/10.1007/s40593-018-0165-4

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior*, *39*, 356–367. https://doi.org/10.1016/j.chb.2014.07.018

Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., & Price, M. J. (2019). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system? *Learning and Instruction*, *5*(April), 101200. https://doi.org/10.1016/j.learninstruc.2019.04.001

Taub, M., Mudrick, N. V., & Azevedo, R. (2018). Strategies for designing advanced learning technologies to foster self-regulated learning. *Strategies for Deep Learning with Digital Technology: Theories and Practices in Education*, pp. 137–169.

Todd, R. W. (2006). Continuing change after the innovation. *System*, *34*(1), 1–14.

https://doi.org/10.1016/j.system.2005.09.002

Trevors, G., Duffy, M., & Azevedo, R. (2014). Note-taking within MetaTutor: interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. *Educational Technology Research and Development*, *62*(5), 507–528. https://doi.org/10.1007/s11423-014-9343-8

Tsai, Y. L., & Tsai, C. C. (2018). Digital game-based second-language vocabulary learning and conditions of research designs: A meta-analysis study. *Computers and Education*, *125*(August 2017), 345–357. https://doi.org/10.1016/j.compedu.2018.06.020

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69. https://doi.org/10.1023/A:1005204207369

VanLehn, K. (1988). Toward a Theory of Impasse-Driven Learning. In M. M. Sebrechts, G. Fischer, & P. M. Fischer (Eds.), *Learning Issues for Intelligent Tutoring Systems* (pp. 19–41). https://doi.org/10.1007/978-1-4684-6350-7_2

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, *31*(1), 3–62. https://doi.org/10.1080/03640210709336984

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why Do Only Some Events Cause Learning During Human Tutoring? *Cognition and Instruction*, *21*(3), 209–249. https://doi.org/10.1207/S1532690XCI2103_01

Vilches, M. L. C. (2003). Task-based Language Teaching: The Case of EN 10. *RELC Journal*, *34*(1), 82–99. Retrieved from https://doi.org/10.1177/003368820303400106

Wang, A. (2015). Facilitating participation: Teacher roles in a multiuser virtual learning environment. *Language, Learning and Technology*, *19*(2), 156–176. Retrieved from http://llt.msu.edu/issues/june2015/wang.pdf%0AJune

Wang, C. X., Calandra, B., Hibbard, S. T., & McDowell Lefaiver, M. L. (2012). Learning effects of an experimental EFL program in Second Life. *Educational Technology Research and Development*, *60*(5), 943–961. https://doi.org/10.1007/s11423-012-9259-0

Wang, Y. (2013). Incidental Vocabulary Learning Through Extensive Reading: A Case of Lower-level EFL Taiwanese Learners. *The Journal of AsiaTEFL*, *10*(3), 59–80. Retrieved from https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/2rsddf/cdi_proquest_journals_2266416058

Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE—Virtual Immersive Language Learning and Gaming Environment: Immersion and presence. *British Journal of*

Educational Technology, *48*(2), 431–450. https://doi.org/10.1111/bjet.12388

Wehner, A. K., Gump, A. W., & Downey, S. (2011). The effects of Second Life on the motivation of undergraduate students learning a foreign language. *Computer Assisted Language Learning*, *24*(3), 277–289. https://doi.org/10.1080/09588221.2010.551757

Wichadee, S. (2018). Significant predictors for effectiveness of blended learning in a language course. *JALT CALL Journal*, *14*(1), 25–42. https://doi.org/10.29140/jaltcall.v14n1.222

Wilkins, D. A. (1972). *Linguistics in Language Teaching* (Print). London: Edward Arnold.

Winn, W. (1993). A Conceptual Basis for Educational Applications of Virtual Reality. Retrieved January 6, 2021, from http://www.hitl.washington.edu/research/learning_center/winn/winn-paper.html~

Winne, P. H. (2015). Self-Regulated Learning. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (2nd ed., pp. 535–540). https://doi.org/10.1016/B978-0-08-097086-8.25091-5

Winne, P. H. (2017a). Cognition and Metacognition within Self-regulated Learning. In P. A. Alexander, D. H. Schunk, J. A. Greene, & P. H. Winne (Eds.), *Handbook of Self-Regulation of Learning and Performance Cognition and Metacognition within Self-Regulated Learning* (pp. 36–48). https://doi.org/10.4324/9781315697048.ch3

Winne, P. H. (2017b). The trajectory of scholarship about self-regulated learning. *Teachers College Record*, *119*(13), 1–16. Retrieved from https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/2rsddf/cdi_proquest_journals_1929038873

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice*. New York & London: Routledge.

Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. J. Zimmerman (Eds.), *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (pp. 297–314). https://doi.org/10.4324/9780203831076

Winne, P. H., & Perry, N. E. (2005). Measuring self-regulated learning. In M Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). Retrieved from http://ebookcentral.proquest.com

Yamada, M., Goda, Y., Matsuda, T., Saito, Y., Kato, H., & Miyagawa, H. (2016). How does self-regulated learning relate to active procrastination and other learning behaviors?

Journal of Computing in Higher Education, *28*(3), 326–343. https://doi.org/10.1007/s12528-016-9118-9

Yoshii, M., & Flaitz, J. (2002). Second Language Incidental Vocabulary Retention : The Effect of Text and Picture Annotation Types. *CALICO*, *20*(1), 33–58. Retrieved from http://www.jstor.com/stable/24149608

Zheng, D., Young, M. F., Brewer, R. A., & Wagner, M. (2009). Attitude and self-efficacy change: English language learning in virtual worlds. *CALICO Journal*, *27*(1), 205–231. Retrieved from http://www.jstor.org/stable/calicojournal.27.1.205

Zimmerman, B. J. (2005). Attaining Self-regulation: A social cognitive perspective. In M Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Retrieved from https://ebookcentral-proquest-com.ezproxy.library.sydney.edu.au

Appendix 1 Pre-test and post-test

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Pre-test and Post-test – written assessment

Directions.
There are two sections of this test, Section A will measure your competence in describing an outdoor a place. Section B will measure your skill in describing the outdoor place. Please read the directions carefully before you do each section.

Section A.
The following questions will be based on the picture.

1. There are multiple tall, dark green palm trees ……. (1) of the image. On the right-hand side, ……. (2) a few shorter light green palm trees.

2. There are many clustered light and dark green trees beside the ocean's edge. They stand ……. (3) some of the tall, dark green palm trees.

3. Blissful blue skies fill the scene. Often little ……. (4) clouds are passing through the sky.

4. A few white fluffy clouds gently ……. (5) through the bright blue sky.

5. A brown coloured ……. (6) detailed with two blue vertical stripes in its centre, ……. (7) upright against the stilt house. There is also another surfboard ……. (7) upright leaning against a chair ……. (8) the ocean. This second surfboard is only visible from the back.

6. The light blue ocean calmly ……. (9) during the daytime. I can see and hear the melodic ocean waves crashing against the rocks and the shore.

7. Growing on large brown rocks, is moss and green plants with many long and thin ……. (10) leaves.

8. A green plant with ……. (11) flowers through it appears on top of a brown rock, and on the ……. (12).

9. Beautiful yellow beach sand is ……. (13) the shore.

10. Light brown ……. (14) are positioned on the sand beside the water's edge. They are facing ……. (15) to the shore's edge. These two chairs ……. (16) underneath a brown straw shelter.

11. A brown straw ……. (16) in the shape of a circular prism is positioned on the sand beside the ocean. It ……. (17) shade to the two brown beach chairs underneath it.

12. A brown ……. (18) house made of wood sits over the ocean's edge. Its windows are wide open.

13. A dark green plant with large wide leaves ……. (19) out of a brown rock.

Appendix 2 Spoken Assessment

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Pre-test and Post-test – performance assessment

Direction.
We are going to have a gathering in Second Life. We would like you to recommend a place. Please consider the followings when choosing the place

1. Around 20 people are going to attend the gathering
2. The audience are already very busy in the real world. They need a relaxing place where they can listen to music or nature. They might want to dance with beautiful music.
3. If you are one of the participants, what do you like and what don't you like about the place.
4. you will have to present the place in front of the prospective participants of the gathering.

Students are going to choose from these five pictures. When they have chosen one, they need to copy and paste the link to their second life and use the video to emphasize their description.

1. http://maps.secondlife.com/secondlife/Fire/45/34/22

2. http://maps.secondlife.com/secondlife/Jewels%20Islands/81/63/22

3. http://maps.secondlife.com/secondlife/Flotsam%20Beach/49/165/22

4. http://maps.secondlife.com/secondlife/VIRTLANTIS%20Community/13/242/21

5. http://maps.secondlife.com/secondlife/La%20Cumbrita/210/148/21

Appendix 3 Open-ended pre-survey

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Survey on students' background

Instructions: Please complete the following questions to reflect your opinions as

accurately as possible and to answer factual questions to the best of your knowledge.

Your information will be kept strictly confidential.

Name:

Origin :

English Tests results:

Have you lived in English speaking countries before? How long have you lived there?

How long have you learned English language? Where did you learn the English language before? Have you played Second Life before?

If yes, how often do you play Second Life?

'

Appendix 4 Exit Survey

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Exit Questionnaire

Instructions: Please complete the following questions to reflect your opinions as accurately as possible and to answer factual questions to the best of your knowledge.

Your information will be kept strictly confidential.

1. How did you like learning in the research?
2. What seems to be helpful from the learning experience you just experienced?
3. What suggestion might you have to improve it?
4. How did you find the consolidation helpful with your speaking development?

Appendix 5 Raw data on written assessment

Name	1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		Total
muthiafhm	in the left corner	0	there are	1	right next to	0	thin fluffy	0	pass	1	surfboard	1	on	0	wave	0	under the	0	rose, sand	0	near	0	under the umbrella	0	inside	0	umbrella	0	is	0	above	0	3
codelight	in the left of picture and in the right picture near the rock	0	a dark green trees	0	right beside the palm trees	0	beautiful	0	dancing	0	surfboard	1	lay on	0	sing	0	moss	0	red, sand	1	beautify	0	house/hotel	0	under	1	brown straw shelter	0	leafs	0	on	0	3
avisena	in left	0	I can see	0	in right	0	in under	0	in above	0	near by	0	inside	0	weather	0	rock	0	red, white sand	0	nice	0	in room	0	near by	0	umberella	0	tree	0	tree	0	0
mulindafir	in the left	0	there are	1	together	0	of	0	covering	0	surfboard	1	in	0	move	0	beautiful	0	pot, sand	0	in front of	0	of straw shelter	0	on	0	umbrella	0	which	0	on	0	2
abelynteresa	at the corner of the upper left	0	beside the dark palm trees	0	next to	1	fluffy	1	pass	1	surf boarding	0	in front of	0	move to the edge of the beach	0	sharp	0	beautiful, sand	0	spread	1	at the edge of the beach	0	in front of	0	umbrella	0	shelter	0	on top	0	4
saputrifirda	left	0	in front of	0		0	under	0	top	0		0	in front of	0	sounds	0	around	0	bottom	0	beside	0	on top	0	between	0	umbrella	0		0	on the top	0	0
namdhanitaufik101	on the left side	1	there are	1	on	0	on the sky is	0	on the sky	0	of surfboard	0	beside the	0	waves	0	leaf	0	bit, sand	0	so soft all over	0	on the house	0		0	beach umbrella	0	stone	0	on	0	2
sumiyati	on the left side	1	there is	0	on the left side	0	fluffy	1	passing	0	surfing board	1	on the	0		0	green	0	wild , sand	0	near	0	hut	0	in front of	0	shelter / hut	0	which	0	on	0	3
yusoooo	green parnis	0	oceans	0	fill the scene	0	gently	0	colored	0	surfboard	1	small	0	enjoyed	0	for	0	people way	0	white	0	scence	0	fill the scence	0	surfboard	0	green flower	0	colord	0	1
deviarl	coconut trees	0	palm trees	0	together	0	fluffy	1	with	0	stage	0	stand up in the backside of a chair	0	scenery	0	with	0	stone, beach	0	sun	0	stage	0	umbrella	0	umbrella	0	which	0	garden	0	1
entinisa04	On the left-hand side	0	a few short dark green palm trees	0	a side	0	fluffy clouds	0	passing	0	surf board	0	standing up line -	0	purpose	0	brushal legs	0	with, scene	0		0	beaches	0		0		0		0		0	0
obsidian09	on the left side	1	there are	1	with	0	fluffy	1	coloring	0	surfing board	1	on	0	come	0	thinly	0	beautiful, house	0	fill	0	shore	0	with	0	umbrella	0	house	1	on top	0	5
TheyCallMeIrvan	on my left	0	there are	1	beside of	0	thin	0	show	0	board surf	0	in front of	0		0	green	0	pot, beach sand	0		0	door	0	bottom of	0	umbrella	0	door	0	on top of	0	1
AastKiryuu	Left	0	there are	1	taller	0	cleany	0	pass around	0	board	0	big	0	wave	0	smooth	0	valvate, sand	0	lighting	0	house	0	on	0	umbrella	0	board	0	on	0	1
KrismonikaD	On the left side	1	There are	1	On	0	Beautiful	0	By	0	Surboard	0	beside	0	burn	0	A lot of	0	cute, wood	0	edge	0	chair	0	beside	0	umbrella	0	house	1	on	0	3
jessieparker29	lining near the water	0	there are	1	beside	1	fluffy	1	floating pass	0	surf board	0	on	0		0	sharp	0	red, of the sands	0	filled	0		0	in front of	0	straw shelter	1		0	on top of	0	4
firafrd	garden or jungle	0	trees	0	cloud	0	surfing	0		0	stone	0	sand	0	smoke	0	plant	0	plant	0	water	0	beach	0	beach	0	umberella	0	cottage	0	plant	0	0
natuaini	above	0	beautiful	0	around	0	fluffy	1	are passing	0	surfboard	1	in front of	0	beautiful	0	small	0	sand	1	around	0	inside	0	in front of	0	umbrella	0	building	0	above	0	3
rumayshoo	up	0	flower to	0	up to	0	cloud	0	walk	0	house	0	sit down	0	talking	0	stone	0	change post	0	sound	0	umbrella	0	under sit down in the chair	0	sceatboard	0	trees, world	0	stand up	0	0
MasSeto	Beside	0	there is	0	between	0	blue	0	dissapear	0		0	on	0	gone	0		0	beautiful, sands	0	near	0		0	upon	0	circle	0	square	0	above	0	0
ulpacan	on the left side of the image	0	there are	1	on the right side	0	blissful	0	finally	0	of surfing board	0	The right side of a chair facing the ocean.	0	move	0	small	0	red, sand	1	fill	0	stand	0	underneath	1	straw shelter	1	stairs	0	above	0	4
yasmend	green	0	leaf	0	3 green	0	bird	0	the sky firs so much but the ending is used up	0	borwn	0	just borwn	0	blue	0	growing	0	black colour	0	deseret	0	leaf	0	the chairs	0	the umbrella	0	roof	0	the leaf	0	0
ulfaa11	on	0	ficture	0	on the	0	and blue	0	beautiful	0	stairs	0	to	0	ombak	0	for	0	house	0	nice	0		0	place	0	and green	0	is	0	with	0	0
somiamanis	on left	0	there is	0	besided	0	color	0	air	0	house	0	beside	0	move	0	small	0	small	0	move	0	space	0	under	1	umbrella	0	house	1	on	0	2
fnptri	above	0	dark green	0	on the left	0	slowly	0	bird	0	the sand	0	in front of	0	faster	0	soft	0	soft, flower	0	waves	0	the chair	0	behind	0	umbrella	0	roof	0	above	0	0
ariesapril	most	0	appear	0	by	0	fluffy	1	drift	0	the surfboard	0	on	0	against	0	large wide	0	8.beautiful 9.sand	0	spread out	0	staircase	0	between	0	umbrella	0	resort	0	on	0	1
Alfiyyahnurazizah	tall	0	trees	0	besides	0	beautiful view,	0	and sound of birds chirp, sound of water	0	flowers	0	and	0	hear	0	look	0	beautiful. sand	0	wide	0	house	0	in the middle	0	straw	0	house	1	on	0	1
Lailaam	left	0	in	0	right	0	beautiful	0	is	0	surfboard	1	on	0	can	0	good	0	beautiful	0	belongs	0	wood	0	in	0	umbrella	0	house	1	on	0	2
kullaut	left	0	many trees	0	on	0	bautiful	0	fly	0	surfboard	1	behind	0	fly	0	big	0	beautiful	0	in front of	0	make	0	under	1	umbrella	0	hut	0	on	0	2
AyuLestariKalinda	left side	0	in	0	on	0	beautiful	0	fly	0	surfboard	1	on	0	when	0	big	0	beautiful, sand	0	in front of	0	hut	0	on	0	umbrella	0	hut	0	on	0	1

Questions

1. There are multiple tall, dark green palm trees ..... (1-preposition) of the image.
2. On the right-hand side, ... (2-NV) a few shorter light green palm trees.
3. There are many clustered light and dark green trees beside the ocean's edge. They stand ..... (2-preposition) some of the tall, dark green palm trees.
4. Blissful blue skies fill the scene. Often little white ..... (4-adj) clouds are passing through the sky.
5. A few white fluffy clouds gently ..... (4-V) through the bright blue sky.
6. A brown colored ..... (5-N) detailed with two blue vertical stripes in its center, is stood upright against the stilt house.
7. There is also another surfboard leaning ..... (5-preposition) a chair facing the ocean. This second surfboard is only visible from the back.
8. The light blue ocean calmly .....(6-V) during the daytime. I can see and hear the melodic ocean waves crashing against the rocks and the shore.
9. Growing on large brown rocks, is moss and green plants with many long and thin ...../(7-Adj) leaves.
10. A green plant with little ..... (8-adj) flowers through it appears on top of a brown rock, and on the ..... (9-N).
11. Beautiful yellow beach sand is ..... (9-V) the shore.
12. Light brown recliner..... (12-N) are positioned on the sand beside the water's edge.
13. The chairs are facing parallel to the shore's edge. These two chairs are positioned .....(10-preposition) a brown straw shelter.
14. A brown ...../(11-N) in the shape of a circular prism is positioned on the sand beside the ocean. It provides shade to the two brown beach chairs underneath it.
15. A brown stilt ...../(12-N) made of wood sits over the ocean's edge. Its windows are wide open
16. A dark green plant with large wide leaves has grown ..... (13-prep) a brown rock

Appendix 6 Example of oral test transcription and coding

Media File: Post-test Oral
Subject: Yasmend
Compliancy: non-compliant (no task 2)
Group: PF
Move to: -

	Timespan	**Content**
1	00.18.48.05 – 00.18.56.19	Okay, um, in here, I'm in a beach (siting on the rock)
2	00.18.57.09 – 00.19.11.08	In beach, there are, um, ….. green, ….. you can sit down in here and you can look
3	00.19.12.23 – 00.19.29.12	In front of and, can we turn around, In here, …. Um, the water is blue like a sky. But in the sky are like, .. a blue
4	00.19.30.24 – 00.19.44.15	Um, a black, . . and after there, there are, … a bridge, . there are bright
5	00.19.47.10 – 00.19.49.17	And then, …..
6	00.19.54.24 – 00.19.56.08	There are a coconut. (zooming out)
7	00.19.59.00 – 00.20.07.02	And I can um, living in, …. Um, a *brick?
8	00.20.09.21 – 00.20.15.18	In here, there are um, ….. umbrella like umbrella
9	00.20.17.18 – 00.20.26.25	(rotating the camera) and then, ….
10	00.20.28.07 – 00.20.41.00	In front of me, ….. there are a coconut (rotating the camera view)
11	00.20.47.29 – 00.20.52.22	In the beach, there are a water.
12	00.20.55.03 – 00.20.57.14	So many water (continue rotating the camera)
13	00.21.01.22 – 00.21.10.00	And, I think have done. And thank you. …
14	00.21.10.23 -	(calling the volunteer guide)

15

16	00.21.14.11 – 00.21.28.06	Halo, (volunteer:Halo) I have done, (Volunteer: okay thank you, um, just wait)

17	00.21.57.11 – 00.21.59.23	I must wait? (Talking to volunteer)
18	00.22.05.24 – 00.22.11.11	(Volunteer: Yes. Um, .. where are you?) I'm sit down.
19	00.22.12.24 – 00.22.17.08	(Volunteer: uh no, I mean in secondlife)
20	00.22.24.20 – 00.22.34.01	(Volunteer: could you describe a bit the place?, oh, there you are, okay)
21	00.22.39.27 – 00.22.46.23	(Volunteer: what do you think about the place?) I think the place um, so beautiful.
22	00.22.47.14 – 00.22.55.25	(Volunteer: uh huh!) and, um, … *sockle? Right. In this place.
23	00.22.57.14 – 00.23.02.09	I think um, ve-, very very nice.
24	00.23.13.12 – 00.23.19.00	(Volunteer: um, I think a bit noisy here)
25	00.23.26.05 – 00.23.36.20	(Volunteer: what do you think about the tree?) um, the tree, the color is, um, white?
26	00.23.38.15 – 00.24.06.17	(Volunteer: um, it little bit green) no, but I happy for a play a tree, and no, the tree, um, very much. For, um, for example um, tree, a coconut. (Volunteer: coconut tree aha!)
27	00.24.08.00 – 00.24.23.19	And, um, when I, when I sss-, seen, I seen a coconut. Thrown in front of, in front of me.
28	00.24.23.19 – 00.24.27.29	And so many, I'm so excited,
29	00.24.29.25 – 00.24.46.03	(Volunteer: very good) ….. when the origin, um, I, I drink the coconut. But this is um, just imagination
30	00.24.47.05 – 00.24.56.25	(Volunteer: it's okay, very good I think. Have you um, done your post test three?) yes I have done!
	00.24.57.22 – 00.25.07.14	(Volunteer: ok, thank you. Um, did you submit it? All of it?) …. Um,
	00.25.15.08 – 00.25.32.09	Describe some of the videos right? (Volunteer: aah, and then um, you should enter the answer and then um, .. click the submit button.) yes I have done.
	00.25.32.26 – 00.26.02.05	(Volunteer: ok, thank you, thank you very much. I think this is the end of our session today, I do, um, thank you for your corporation. And then hope you will be learn something a good today. And then hope um, we can see face to face someday. Thank you) okay..
	00.26.02.05 – 00.26.11.02	(Volunteer: Good bye, bye!) se you!

AN EXAMPLE OF CODING ON THE SPOKEN DESCRIPTION OF PLACES

Participants	length	messages	correct sentencs	Score		Verbs		adj		prep		things at the beach		Syntaxt	
				vocab	grammar										
jugheaddd (Nvivo)	5:56	9	7	17	6		3		5		2		7		6
(Opening Video from YouTube)															
Okay, um, this is a fire beach. So can-, I can hear um, and I-, I can see and I can hear. Um, ... wait, .. um, ..		1	1			see	1					a fire beach	1	I can see	1
Um, there are a lot of coconut trees, a group of coconut-, a group of tall coconut trees. And also, .. um, I can, (she paus the video)		1	1					tall	1			coconut tree	1	There are	1
I think is a-, *comfort?. Comfort place to get in run with about twenty people. Because everybody, everyone needs relaxing and I think it's a best place to relax.		1	1					comfortable	1					It is AdjN	1
Um, at this place we can also, um, .. dancing with our friends		1				dance	1							S can V	1
And then, I can, .. you can see, you can also see a beautiful sky above your head. .. and the clou- and, ... and the cloud passing through sky. Um, ...		1	1					beautiful	1	above	1	sky	1	N ving	1
I can see a little plants, and red flowers. It's so beautiful red flowers.		1	1			see		little	1			plants, flowers	2		
And also, ... um, .. on the left side, I can see um, .. palm trees,		1	1							on the left side	1	palm trees	1		
And, ... But I can't see surf- surfboard here,.		1	1			play	1					surf board	1		
And you can also play with the yellow sand here.		1	1					yellow	1					sand	1
(taking off her headphone and put it back) (the description is end in here)															
		9	8				3		5		2		7		6

Appendix 7 Lesson plan of the consolidation

THE UNIVERSITY OF SYDNEY
ABN 15 211 513 464

Sydney School of Education and Social Work

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Lesson Plan for TBLT/DI Group

Teacher: Puji Rahayu
Topic: Describing Places (Outdoor)
Focus: Vocabulary and Grammar for Describing Places (Outdoor)

CCSS/CLGs/SC Assessment Limits/Standards: *(What are the skills being taught? Which standards are being specifically addressed in this lesson?)*	**Agenda: (What is the snapshot of my class flow?)**
- Knowledge of a range of vocabularies for describing places (noun, adjectives, preposition of places) - Grammatical knowledge of describing places (there is/are, I can see/hear, Something stands/sits) - Oral description of a virtually outdoor place in SL	1. Watching interactive video 2. Vocabulary and grammatical quiz 3. Challenge Task 1 describing a virtually outdoor place 4. Challenge task 2 describing a virtually indoor place 5. Challenge task 3 describing a virtually outdoor place
Lesson Objective: *(What will my students KNOW by the end of the lesson? What will they DO to learn it?)* - Students are able to use a range of vocabularies for describing places (noun, adjectives, preposition of places) with intelligible pronunciation. - Students are able to use grammatical knowledge of describing places (there is/are, I can see/hear, Something stands/sits) intelligibly. - Students are able to use describe virtually outdoor places in SL with quite wide range of vocabularies and grammar	

TIME	INSTRUCTIONAL SEQUENCE	FORMATIVE ASSESSMENT
		Note: A variety of formative assessments should be used at key points throughout the lesson.
45 min	**Consolidation** • Students will watch a video on describing a place they just described in Task 1 A video viewing will take place for 5 minutes. • Students and teacher will have a discussion on what to cover in a description Okay, could you please mention aspects for describing places? Name of the place, things in the place and their position, interesting things in the place, what to do in the place Yes, we need to include answers to the following questions - What is the name of the place? - Where is the place? - How can you describe the place itself? - What are the key facilities and activities to do in the place - how do you feel about the place? • Students will listen to a lecture on the vocabulary used in describing places, both the meaning and pronunciation Now, we'll see some pictures and I would like you to say the words.	

Noun/ things	Prepositions	Adjectives	Adverbs	Verb
Palm trees	On the right side	Melodic	Gently	Stand
Ocean's edge	On the left side	Fluffy	Calmly	Sit
Sky	Against	Blissful		Pass
Clouds	On top of	Small/ big		is/am/are
Surf board	Underneath	Long/tall/short		Ripple
Stilt House	Over	Small/ big/ large/ little		
Ocean	Out of	Moss		
Waves		Moss		
Rocks		Bristle-like		
Shore		Stilt		
Plants				

		Leaves				
		Flowers				
		Sand				
		Moss				
		Recliner chair				
		Edge				
		Shore's edge				
		Straw shelter				
		Shade				
		Stilt house				

Which nouns can go with which adjectives?
Fluffy clouds
Stilt house
Tall palm trees
Melodic waves
Well done!
Look at the picture!
We call this plant "a coconut tree". We can say "a group of coconut trees".
Some adjectives that can go with coconut tree are tall, skinny, green.
So we can say a group of tall skinny coconut trees.

- Students will do a vocabulary quiz on line.
  Links provided (5 minutes)
- Students will listen to a lecture on the grammatical structure used in describing places
  Now, I would like you to make sentences using those words based on the video.
  There is a group of tall skinny coconut trees on the left side of my place.
  I can see a surf-board standing in front of the stilt house.
  Excellent!
- Students will listen to a lecture on the grammatical structure used in describing places
  We can use
  There is/ are …… on the left/ right/ in front of/ beside ……
  I can see/hear …… on the left/ right/ in front of/ beside….
  The melodic waves at the edge of the sea are very beautiful and releasing.
- Students will do a quiz on the use of vocabulary and grammar in describing places.
  Links provided (5 minutes)

10 min	**Challenge1/ Task 1** • Students will describe Virtlantis1, a beach in Second Life, as a place for a social gathering. Instruction will be delivered by a virtual conversation companion.	
10 min	**Challenge 2/ Task 2** • Students will describe Virtlantis2, a meeting room in Second Life, as a place for a conference meeting. Instruction will be delivered by a virtual conversation companion.	
10 min	**Challenge 3/ Task 3** • Students will choose a place from the list and describe it as the place for ….. Instruction will be delivered by a virtual conversation companion.	
75		

Screen shot of the place for challenge 1	

Screen shot of the place for challenge 2	
Screen shot of the place for challenge 3/ Post-test	

Appendix 8 Information pack statement for conversation companion

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Informasi tentang ketentuan-ketentun menjadi relawan
dalam penelitian MUVE-based English language learning

Dear Teman-teman,

Terimakasih sudah menyampaikan ketertarikan teman-teman menjadi volunteer dalam penelitian saya. Saya akan mengadakan seleksi kecil untuk teman-teman yang bisa menjadi volunteer.

Informasi tentang *Volunteering* (10 - 15 orang)

1. Teman-teman akan bekerja berdasarkan jadual yang akan saya berikan (sesuai dengan ketersediaan waktu partisipan) dalam rentang waktu tanggal 17 - 29 April 2018
2. Dalam kegiatan *paid voluntry* ini teman-teman akan menemani beberapa partisipan dalam belajar Bahasa Inggris di *SecondLife*
3. Semua pekerjaan akan dilaksanakan di *SecondLife*
4. Jumlah jam kerja akan dihitung ketika saya sudah menyatakan bahwa teman-teman bisa terlibat dalam penelitian saya
5. Setelah menerima informasi dari saya kemudian teman-teman berubah pikiran, dipersilahkan "leave" group ini. Anda tidak perlu merasa tidak enak dan tidak perlu minta ijin (meskipun minta ijin juga boleh2 saja).
6. Apabila teman-teman ada pertanyaan, silakan di post di group ini. Saya akan menjawab pertanyaan2 tersebut juga di group ini.
7. Pengumuman berikutnya akan dilakukan dengan menggunakan nama *SecondLife* yang teman-teman buat.

Langkah-langkah untuk menjadi volunteer

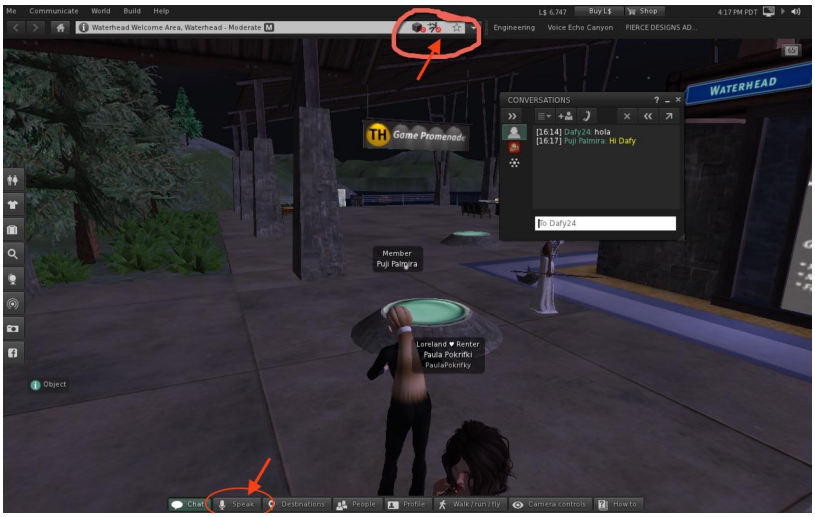
Saya menyarankan teman-teman untuk melakukan langkah2 selanjutnya:

1. Membuat akun di secondlife.com (karena aktifitas kelas akan dilakukan di Secondlife)
2. Mengekplorasi *SecondLife* sampai teman-teman bisa memastikan bahwa *hardware* teman-teman stabil untuk *voice chat*. Teman-teman bisa memaksimalkan fungsi *search* untuk mencari tempat-tempat yang banyak dikunjungi orang-orang.
3. Melakukan voice-chat dengan saya atau asisten penelitian saya untuk mengetahui apakan sistem teman-teman oke untuk voice chat di *SecondLife* (Saya akan memilih yang paling sedikit bantuan teknis yang harus saya berikan - mohon maaf ini alasan practical saja). Voice chat insyaAllah akan dilaksanakan pad Senin 16 April 2018 antara jam 4 - 8 pm WIB.

Voice chat

Saya yakin bagi teman-teman yang sudah biasa main game online sudah sangat mengetahui tentang fungsi voice chat. Dalam *SecondLife*, ada 2 macam voice chat:

1. *Nearby chat.* Fungsi ini untuk berkomunikasi suara dengan orang-orang sekitar kita. Dengan demikian, kalau kita agak berjauhan dengan pembicara kita akan kurang mendengarkan suaranya. *Nearby chat* bisa dilakukan dengan menekan *"speak"* di kiri bawah dengan syarat tidak ada simbol "*no-voice"* di atas (orange).
2. *VoiceCall.* Fitur ini digunakan untuk berbicara dengan orang (orang-orang) tertentu yang tidak harus satu *frame* dengan kita. Dengan *VoiceCall* ini suara kita hanya akan terdengar oleh orang-orang yang kita telepon.



Links tentang *SecondLife*

https://strawberrysingh.com/2016/08/31/introduction-second-life/

Updates

17 April 2018

Bagi yang tadi malam belum sempat bertemu saya, silakan online siang-sore ini jam 2-4 WIB.

Jam 4 sore, saya akan mengirimkan pesan di akun SL teman-teman tantang hasil dari pertemuan kmaren.

Thanks

Tempat pertemuan untuk tanggal 17 April 2018 jam 14.00 di

http://maps.secondlife.com/secondlife/Builders%20Brewery/112/157/21

19 April 2018

- Saya sudah mengundang teman-teman yang di sebuah dokumen untuk pedoman saat menemani peserta. Bagi yang sudah chat dengan saya dan tidak ada masalah teknis tapi belum dapat dokumennya, silakan japri di nomer WA saya.
- Teman-teman silakan mampir ke SecondLIfe Indonesia Headquarter, naik ke atas pakai fungsi teleporting, dan mencoba memutar video lewat youtube player yang ada di lantai atas. Sila masuk lewat link ini http://maps.secondlife.com/secondlife/Mangakino/37/37/22
- Silakan eksplore SecondLife more supaya semakin mahir.
- Mohon maaf kalau ada kekurangan dalam seleksi ini.

19 April 2018

Link berikut berisi tentang tips untuk memperbaiki performa voice chat kita menggunakan microfon di Second Life.

https://www.sl-inspiration.com/2017/10/how-to-improve-quality-of-sound-from.html

Appendix 9 Information statement for participants

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Productive Failure in Multi-user Virtual Environments (MUVE)-based English as a Foreign Language (EFL) Activities: Students' Performance in oral vocabulary and grammar

You are invited to take part in a research project. This document contains important information about the project and the roles of potential participants. Please ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. *Thank you for reading this.*

INFORMATION STATEMENT

(1) What is the study about?

This project is looking at the development of oral proficiency in online-based English as a foreign language activities among intermediate students in Indonesia, living in Yogyakarta. Two different pedagogical strategies will be used, productive failure and task-based language teaching/ direct instruction. Participants will be randomly divided into two groups. Both groups will do some activities of describing virtual places in Second Life, a multi-user virtual environment. One group will describe the places prior to the presentation session while the other group will describe the place after the presentation. Participants will have to do pre-and post-test activities of describing places in Second Life. For the purpose of data collection, students' screen will be video recorded while they are doing online activities. The online activities will be video recorded as well for the purpose of dissemination.

(2) Who is running the study?

The study is being carried out by the following researchers:

- Michael J Jacobson, PhD., Sydney School of Education and Social Work, Research Professor and Chair of Education, as the Chief Investigator,
- David J Hirsh, PhD., Sydney School of Education and Social Work, Associate Professor in TESOL, as a co- investigator, and
- Ms. Puji Rahayu, Sydney School of Education and Social Work, PhD student (Co-Investigator)

(3) What are we asking you to do?

We are asking you for two visits to the language laboratory at English Education Department Universitas Islam Indonesia. You will do online activities in Second Life where you can develop your oral proficiency on describing virtual places.

You will be asked:

1. On the first visit;
   - to join a training on participating online classes in *Second Life* approximately for one-hour face to face at the language laboratory.
   - to fill in a ten-minute online survey about your English language background in Moodle attached to *Second Life*. This survey will be about i) your general information, ii) your experience of playing games and participating in online classes, and iii) your exposure to the target language communities.
   - to log into Second Life for an online pre- and post-test sessions, which will record your speaking proficiency and speaking skills on describing places. This will include a cloze procedure test and a performance-based oral proficiency test approximately for thirty minutes. The pre- and post-tests in Second Life will be screen-captured in video format.
2. On the second visit;
   - to do activities in *Second life* based on the assigned group (control or experimental group). You will describe two virtual places in Second Life and listen to a consolidation in an interactive video format. The proctor will help you with technical challenges you might be facing during your online activities. This session will be screen-captured in video format via students' computer and video-recorded to assist with the data collection and analysis. This activity will take approximately 90 minutes and be screen-captured in video format.
   - fill in a ten-minute online survey in Moodle attached to *Second Life* on your experience after taking part in the study.

(4) How much time will the study take?

It will take approximately three hours twenty minutes or 200 minutes. You will receive a timetable of the sessions and will be asked to choose one preferred time slot. A video on how to use Second Life will be shared prior to the training.

(5) Who can take part in the study?

We are looking for second semester English department students from Universitas Islam Indonesia, Universitas Negeri Yogyakarta, and Universitas Muhammadiyan Yogyakarta, Indonesia.

(6) Can we withdraw from the study?

Participation in this study is completely voluntary and you are free to withdraw from the project at any time. Your decision to withdraw will not affect your current or future relationship with your status at your university, or the researchers or anyone else at The University of Sydney. Your data-related activities will not be included in the analysis upon your withdrawal.

(7) Are there any risks or costs associated with being in the study?

There is no risks or costs associated with the study. The places in the virtual environment will be selected in order that students will not get affected by possible misconducts.

(8) Are there any benefits associated with being in the study?

By participating in this study, you will be given opportunities to develop both your speaking proficiency and speaking skills related to describing places.

Your participation in the project will improve your skills in attending a Second-life based English class. This skill can be applied in different learning opportunities provided that Second Life has a number of language classes and professional groups.

To express our appreciation of your contribution, we will provide a Certificate of Participation and a voucher worth AUD20.

(9) What will happen to information that is collected during the study?

All data collected in this study will be kept strictly confidential, except as required by law. All data will be stored electronically in a secure network drive provided by the University of Sydney. Hard copy materials will be stored in a locked cabinet in the office of the chief investigator. Only the approved researchers listed above will have access to the data.

The results of this study may be published in academic journals or used for teaching purposes. The results may also be presented at academic meetings or in talks at academic institutions. Results will always be presented in such a way that data from individual volunteers cannot be identified. For example, the participants' face will be blurred whenever videos are used, or pseudonym will be used instead of the participants' names.

(10) Can I tell other people about the study?

Yes, you are welcome to tell other people about the study.

(11) What if I require further information about the study or my involvement in it?

When you have read this information, Ms Puji Rahayu will be available to discuss it with you further and answer any questions you may have. If you would like to know more at any stage during the study, please feel free to contact her. Ms. Puji Rahayu is available on prah0535@sydney.edu.au, +62274898444 ext. 2107, or +61414886778 (WhatsAp).

(12) Will I be told the results of the study?

You have a right to receive feedback about the overall results of this study. You can tell us that you wish to receive feedback by ticking the relevant box on the consent form. This feedback will be in the form of a one-page summary. Feedback will be available at *Info Academik* section at http://pbi.uii.ac.id after approximately 3 months from the data collection.

(13) What if I have a complaint or any concerns?

Research involving humans in Australia is reviewed by an independent group of people called a Human Research Ethics Committee (HREC). The ethical aspects of this study have been approved by the HREC of the University of Sydney (Approved on DD MM 2017). As part of this process, we have agreed to carry out the study according to the National Statement on Ethical Conduct in Human Research (2007). This statement has been developed to protect people who agree to take part in research studies.

If you are concerned about the way this study is being conducted or wish to make a complaint to someone independent from the study, please contact the university using the details outlined below. Please quote the study title and protocol number.

The Manager, Ethics Administration, University of Sydney:

- **Telephone:** +61 2 8627 8176
- **Email:** ro.humanethics@sydney.edu.au
- **Fax:** +61 2 8627 8177 (Facsimile)

The Dean of Faculty of Psychology and Socio-cultural Sciences, Universitas Islam Indonesia:

- **Telephone:** +62 274 898444 ext. 2106
- **Email:** fpisb@uii.ac.id
- **Fax:** +62 274 898444 ext. 2106

This information sheet is for you to keep

Appendix 10 Information statement for faculties

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Productive Failure in Multi-user Virtual Environments (MUVE)-based English as a Foreign Language (EFL) Activities: Students' Performance in oral vocabulary and grammar

Your students are invited to take part in a research project. This document contains important information about the project and the roles of potential participants. Please ask us if there is anything that is not clear or if you would like more information. *Thank you for reading this.*

INFORMATION STATEMENT

(1) What is the study about?

This project is looking at the development of oral proficiency in online-based English as a foreign language activities among intermediate students in Indonesia, living in Yogyakarta. Two different pedagogical strategies will be used, productive failure and task-based language teaching/ direct instruction. Participants will be randomly divided into two groups. Both groups will do some activities of describing virtual places in Second Life, a multi-user virtual environment. One group will describe the places prior to the presentation session while the other group will describe the place after the presentation. Participants will have to do pre-and post-test activities of describing places in Second Life. For the purpose of data collection, students' screen will be video recorded while they are doing online activities. The online activities will be video recorded as well for the purpose of dissemination.

(2) Who is running the study?

The study is being carried out by the following researchers:

- Michael J Jacobson, PhD., Sydney School of Education and Social Work, Research Professor and Chair of Education, as the Chief Investigator,
- David J Hirsh, PhD., Sydney School of Education and Social Work, Associate Professor in TESOL, as a co- investigator, and
- Ms. Puji Rahayu, Sydney School of Education and Social Work, PhD student (Co-Investigator)

(3) What are procedures of the data collection?
Upon the university approval, the student researcher will:

1. Consult the publicly available schedule to find lecturers teaching "Listening and or Speaking" subjects for semester 2 students;
2. Contact lecturers via publicly available emails to share in their classes about student participation in the project;
3. Upon their approval, the student researcher will come at the end of the class for the information session. Participant Information Statement will be handed to the students at the end of the information session;
4. Students will contact the student researcher if they have questions about the participation;
5. If they wish to participate, they will sign the consent form. The schedule of the Second Life training and the data collection will be shared to the potential participants personally when they hand in the signed consent forms;
6. Students will join the Second Life training and Second Life – based activities based on the preferred schedule at Language Laboratory of English Education Department Universitas Islam Indonesia;

(4) What are we asking your students to do?

We are asking your students for two visits to the language laboratory at English Education Department Universitas Islam Indonesia. They will do online activities in Second Life where they can develop their oral proficiency on describing virtual places.

They will be asked:

1. On the first visit;
   - to join a training on participating online classes in *Second Life* approximately for one-hour face to face at the language laboratory.
   - to fill in a ten-minute online survey about their English language background in Moodle attached to *Second Life*. This survey will be about i) their general information, ii) their experience of playing games and participating in online classes, and iii) their exposure to the target language communities.
   - to log into Second Life for an online pre- and post-test sessions, which will record their speaking proficiency and speaking skills on describing places. This will include a cloze procedure test and a performance-based oral proficiency test approximately for thirty minutes. The pre- and post-tests in Second Life will be screen-captured in video format.
2. On the second visit;
   - to do activities in *Second life* based on the assigned group (control or experimental group). They will describe two virtual places in Second Life and listen to a consolidation in an interactive video format. The proctor will help them with technical challenges they might be facing during the online activities. This session will be screen-captured in video format via students' computer and video-recorded to assist with the data collection and analysis. This activity will take approximately 90 minutes and be screen-captured in video format.
   - fill in a ten-minute online survey in Moodle attached to *Second Life* on their experience after taking part in the study.

(5) How much time will the study take?

It will take approximately three hours twenty minutes or 200 minutes. They will receive a timetable of the sessions and will be asked to choose one preferred time slot. A video on how to use Second Life will be shared prior to the training.

(6) Who can take part in the study?

We are looking for second semester English department students from Universitas Islam Indonesia, Universitas Negeri Yogyakarta, and Universitas Muhammadiyan Yogyakarta,

Indonesia.

(7) Can my students withdraw from the study?

Participation in this study is completely voluntary and your students are free to withdraw from the project at any time. Their decision to withdraw will not affect their current or future relationship with their status at your university, or the researchers or anyone else at The University of Sydney. Their data-related activities will not be included in the analysis upon their withdrawal.

(8) Are there any risks or costs associated with being in the study?

There is no risks or costs associated with the study. The places in the virtual environment will be selected in order that students will not get affected by possible misconducts.

(9) Are there any benefits associated with being in the study?

By participating in this study, your students will be given opportunities to develop both their speaking proficiency and speaking skills related to describing places.

Their participation in the project will improve their skills in attending a Second-life based English class. This skill can be applied in different learning opportunities provided that Second Life has a number of language classes and professional groups.

To express our appreciation of their contribution, we will provide a Certificate of Participation and a voucher worths AUD20.

(10) What will happen to information that is collected during the study?

All data collected in this study will be kept strictly confidential, except as required by law. All data will be stored electronically in a secure network drive provided by the University of Sydney. Hard copy materials will be stored in a locked cabinet in the office of the chief investigator. Only the approved researchers listed above will have access to the data.

The results of this study may be published in academic journals or used for teaching purposes. The results may also be presented at academic meetings or in talks at academic institutions. Results will always be presented in such a way that data from individual volunteers cannot be identified. For example, the participants' face will be blurred whenever videos are used, or pseudonyms will be used instead of the participants' names.

(11) Can I tell other people about the study?

Yes, your students are welcome to tell other people about the study.

(12) What if I require further information about the study or my involvement in it?

When your students have read this information, Ms Puji Rahayu will be available to discuss it with them further and answer any questions they may have. If they would like to know more at any stage during the study, they can contact her. Ms. Puji Rahayu is available on prah0535@sydney.edu.au, +62274898444 ext. 2107, or +61414886778 (WhatsAp).

(13) Will I be told the results of the study?

You and your students have a right to receive feedback about the overall results of this study.

You can tell us that you wish to receive feedback by ticking the relevant box on the consent form. This feedback will be in the form of a one-page summary. Feedback will be available at *Info Academik* section at http://pbi.uii.ac.id after approximately 3 months after the data collection.

(14) What if I have a complaint or any concerns?

Research involving humans in Australia is reviewed by an independent group of people called a Human Research Ethics Committee (HREC). The ethical aspects of this study have been approved by the HREC of the University of Sydney (Approved on DD MM 2017). As part of this process, we have agreed to carry out the study according to the National Statement on Ethical Conduct in Human Research (2007). This statement has been developed to protect people who agree to take part in research studies.

If you are concerned about the way this study is being conducted or wish to make a complaint to someone independent from the study, please contact the university using the details outlined below. Please quote the study title and protocol number.

The Manager, Ethics Administration, University of Sydney:

- **Telephone:** +61 2 8627 8176
- **Email:** ro.humanethics@sydney.edu.au
- **Fax:** +61 2 8627 8177 (Facsimile)

The Dean of Faculty of Psychology and Socio-cultural Sciences, Universitas Islam Indonesia:

- **Telephone:** +62 274 898444 ext. 2106
- **Email:** fpisb@uii.ac.id
- **Fax:** +62 274 898444 ext. 2106

This information sheet is for you to keep

Appendix 11 Consent form

THE UNIVERSITY OF SYDNEY

Sydney School of Education and Social Work

ABN 15 211 513 464

Prof. Michael J. Jacobson, Ph.D.
Professor and Chair of Education

Room 257, Education Building A35
The University of Sydney
NSW 2006 AUSTRALIA
Telephone: +61 2 90367671
Facsimile: +61 2 93512606
Email: michael.jacobson@sydney.edu.au
Web: http://www.sydney.edu.au/

Productive Failure in Multi-user Virtual Environments (MUVE)-based English as a Foreign Language (EFL) Activities: Students' Performance in oral vocabulary and grammar

PARTICIPANT CONSENT FORM

I, ... [PRINT NAME], consent to participate in this research.

In giving my consent I acknowledge that:

1. The purpose of the study, procedures required for the participation, the time, any risks/benefits involved have been explained to me, and any questions I have about the project have been answered to my satisfaction.

2. I have read the Participant Information Statement and have been given the opportunity to discuss the information on my involvement in the project with the researcher/s if I wish to do so.

3. I understand that participation in this study is completely voluntary – I am not under any obligation to consent.

4. I understand that:

   - I am free to withdraw from the project at any time and is free to decline to answer particular questions.

   - While the information gained in this study will be published as explained, my name will not be identified, and individual information will remain confidential.

   - I may ask that the recording/observation be stopped at any time, and I may withdraw at any time from the session without disadvantage.

- My withdrawal will not affect my study and my relationship with the researcher(s) or the University of Sydney now or in the future. My data-related activities will not be included in the analysis upon my withdrawal.

I understand that short extracts of collected data shall be presented at academic conferences. However, these extracts will be carefully edited to mask participants' identities. Also, data will not be shared with anyone other than the approved researchers.

I consent to:

- Fully participating in the online activities YES ☐ NO ☐
- Screen capturing (video) of my Online activities YES ☐ NO ☐
- Video recording of my online activities YES ☐ NO ☐
- Presentation of collected data at academic meetings YES ☐ NO ☐

Would you like to receive feedback about the overall results of this study?

YES ☐ NO ☐

If you answered YES, please refer to the *Informasi Akademik* section of http://pbi.uii.ac.id three months after the data collection.

Would you like to receive information on other research participation opportunities via the address provided above? YES ☐ NO ☐

……..
Signature

……
PRINT name

……
Date

Appendix 12 Samples of exit survey

SL name	How did you like learning in the research?	What is helpful from the learning experience?	What suggestion to improve it?	Is the consolidation helpful for your speaking?	What factor helps you learning?	What hinders you learning?	What did you do while watching the video?	Did the first description help do the second description?	Did you have prior knowledge to do the description?
DI Group									
clearancelvl7	What I like about doing this research is that I can improve my skill in spoken English and also enrich my vocabularies.	What really helpful is that we can interact with an actual native English speakers through the game. It encourages us to speak more.	There are a lot of feature in the game that we dont know yet. We actually can learn using those features while playing but it kinda slows down the process.	I learnt some expressions especially about how to describe an outdoor place.	Like I said earlier, the way it allows me to interact with people around the world helps me to improve my speaking skill.	the internet connection and also the pervs	I tried to memorize the vocabularies and learnt how to pronounce it.	Yes, because I when you try to describe something out loud you unconsciously try to recall your memories about it. That way you can make the memory stays longer.	mostly from the game and Bu Ista's video.

Faridakhuzaimah	bisa bercakap-cakap dengan yang lainnya baik dari orang lokal sampai luar negeri	Yes thats right	untuk mengikuti secondlife	Sangat membantu	faktor pendengaran yaitu ketika kita berbincang dengan yang lainnya dan juga melatih speaking agar lebih lancar	faktor luar yaitu ketika dalam satu tempat semua orang berbicara	Ketika saya melihat video tersebut, saya memerhatikan dengan baik sambil mencermati dan mengingatnya dengan baik, jika ada kalimat yang tidak ketahui saya langsung stop dan mencarinya di google translate.	itu sangat membantu saya untuk melatih kejelian mata saya dalam melihat melihat sekitar sehingga dapat dijelaskan dengan jelas	Saya rasa sudah karena saya sering belajar tersebut pada saat saya sd dan saat les bahasa.
Farrastania23	It was fun to learn english in this way. I mean with game. (secondlife)	The most helpful is you can learn english with instruction from the video and have a guide.	I think thats enough. This is great.	Yes it helps me a lot.	I think that the most helpful things from secondlife is you can have a guide to learn english. I mean if we don't understand something from this learning program, we can just call them and ask.	I think the connection, the voice call doesn't really clear enough.	I think Mrs Ista have a lot of confident to speak in english, she is fluently.	Yes thats help me a lot, because the place was quiet similar with the video on youtube.	yes, I've got it from the video before I click yellow ball such as : Red ball, Purple ball, and and Pink ball.

PF Group									
AtashiDorachii	saya suka karena dalam penelitian ini kita bisa bermain sambil belajar	saat kita melakukan post test, setelah itu kita diberitahu bagaiman cara menjawab yang benar itu seperti apa	mungkin bagi saya dalam penelitian ini diharuskan adanya guide di masing masing individu agar bisa mempermudah saat kita bertanya apabila kita tidak mengetahui suatu hal	iya, sangat membantu	kita bisa melihat hal-hal baru, sambil meneskripsikannya	saat ada orang lain yang berbicara ketika kita sedang berbicara juga dengan orang lain	saya memperhatikan dengan seksama apa yang dilakukan oleh Ibu Ista agar dapat memahami apa yang dijelaskan	ya, dapat menjelaskan . Namun, karena berbeda tempat lagi saya bingung untuk menjelaskan hal baru tersebut	sudah, saya melihatnya di video
Jackhorseman	its fun, playing game is fun	i have the chance to talk with a real person on the game	please improve the instruction and the gameplay experience maybe	the pronounciation, i can learn a lot from it, and also the vocabulary	that second life is a game, and you can learn something quickly when you're enjoying it.	the computer, the lag, and the internet connection	im listening to it	yeah, because i learned a lot of new phrase and vocabulary	from the video

				the video		the signal,the			the prepositions
				explain me		computer			of many things
			maybe added	clearly about the		and	i tried to repated	yes,because	of many things
Leonapatria			more	quiz that i		sometimes	what she said	the video has	and how to
	i can explore		informational	should do	the speaking and the	my second	her and	taught us to	describe
	many places and		video to help the	whether in	chat room helps us	life	remember the	describe in	something in
	i can getting		partipants with	speaking or on	very much to learn	suddenly	material that she	the correct	outdoors places.
	new friends	the video	their quiz	my vocabulary	English	logged out	explain to me	way	From the video

Appendix 13. Ethics approval

Research Integrity & Ethics Administration
Human Research Ethics Committee

Wednesday, 4 April 2018

Prof Michael Joseph Jacobson
School of Education and Social Work Research Operations; Faculty of Arts and Social Sciences
Email: michael.jacobson@sydney.edu.au

Dear Michael Joseph

The University of Sydney Human Research Ethics Committee (HREC) has considered your application.

After consideration of your response to the comments raised your project has been approved.

Approval is granted for a period of four years from **03/04/2018** to **03/04/2022**

Project title: **Productive Failure in Multi-user Virtual Environment (MUVE)-based English as a Foreign Language Learning Activities: Students' Performance in oral vocabulary and grammar**

Project no.: **2018/040**

First Annual Report due: **03/04/2019**

Authorised Personnel: **Jacobson Michael Joseph; Hirsh David; Rahayu Puji;**

Date	Type	Document
14/12/2017	Version 1	Survey on participants background
14/12/2017	Version 1	Exit questionnaire
18/12/2017	Version 1	Pre and Post-test
18/12/2017	Version 1	Lesson Plan
18/12/2017	Version 1	Scoring method
09/03/2018	Version 3	Recruitment Flyer V1 15 Feb 18 - Clean
27/03/2018	Version 3	The clean version of Invitation and proposal UII
27/03/2018	Version 3	The clean version of Invitation UMY
27/03/2018	Version 3	The clean version of Invitation UNY
27/03/2018	Version 5	The clean version of Info Statement for Dean V0 9 March
27/03/2018	Version 5	The clean version of Participant Consent Form V2 27March18
27/03/2018	Version 5	The clean version of Participant Info Statement V2 27March18
28/03/2018	Version 5	The clean version of Email draft for lecturers V1 27March18
28/03/2018	Version 5	The clean version of Info statement for Dean V 1 27March18

Condition/s of Approval

- Research must be conducted according to the approved proposal.
- An annual progress report must be submitted to the Ethics Office on or before the anniversary of approval and on completion of the project.

- You must report as soon as practicable anything that might warrant review of ethical approval of the project including:
  - Serious or unexpected adverse events (which should be reported within 72 hours).
  - Unforeseen events that might affect continued ethical acceptability of the project.

- Any changes to the proposal must be approved prior to their implementation (except where an amendment is undertaken to eliminate *immediate* risk to participants).

- Personnel working on this project must be sufficiently qualified by education, training and experience for their role, or adequately supervised. Changes to personnel must be reported and approved.

- Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, as relevant to this project.

- Data and primary materials must be retained and stored in accordance with the relevant legislation and University guidelines.

- Ethics approval is dependent upon ongoing compliance of the research with the *National Statement on Ethical Conduct in Human Research*, the *Australian Code for the Responsible Conduct of Research*, applicable legal requirements, and with University policies, procedures and governance requirements.

- The Ethics Office may conduct audits on approved projects.

- The Chief Investigator has ultimate responsibility for the conduct of the research and is responsible for ensuring all others involved will conduct the research in accordance with the above.

This letter constitutes ethical approval only.

Please contact the Ethics Office should you require further information or clarification.

Sincerely

[redacted]

Associate Professor Rita Shackel
Chair
Human Research Ethics Committee (HREC 3)

The University of Sydney HRECs are constituted and operate in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2007) and the NHMRC's Australian Code for the Responsible Conduct of Research (2007).

Appendix 14. Observation Sheet

SL name	Group	Length	During the video	realize failure	Finish	subtitle	attention	notes	repeat/ pause	mimick	AHA moment	Memorize	interactive
Fadilla99	PF	41.25	She paused the video and took notes while mimicking the vocabularies mentioned. She turned on the subtitles. She repeated "stilt house" She read the adjectives introduced in the video She described the place when asked.	repeated several times	1	1	1	1	1	1	1	0	1
maladivas	PF	75.00	She paused and took notes on the expressions. She answered the question on aspects of describing places. She mimicked the vocabularies introduced. "owh, stilt house!" and paused and took note for "shelter" while kept sayin the word. She described the place when asked using the N-adj combination that she just learned.	stilt house	1	0	1	1	1	1	1	0	1
AastKiryuu	DI	13.9	watching She mimicked / repeat the words introduced in the video.					0	0	0		0	1
Adhiyaulkhaq	DI	23.4	watched only while playing his pen. stopped in the middle of the video. got called in the middle of hte video. repeated the video once and checked other balls.		straight 1 playing pen 1	1	1	0	1	0		0	0