University of Reading

UNIVERSITY OF READING

Evaluating the Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary in the Classroom: The Case of Oman

Thesis submitted for the degree of

Doctor of Philosophy

Institute of Education

Muna Hussain Muqaibal

January 2020

DEDICATION

Oh Allah, the Almighty! All praise and gratitude be unto You…

ACKNOWLEDGEMENTS

Having thanked Almighty Allah for His blessing in enabling me to reach the end of this research journey, I would like to acknowledge a number of people who have made this journey possible through their support, guidance, assistance and encouragement. Most importantly, my deepest gratitude goes to my first supervisor, Professor Cathy Tissot, for her instructive guidance, which strengthened my writing up of this research and helped me hone my academic skills. I sincerely appreciate her constant support in pushing me to the limit to complete this study. Likewise, I am deeply indebted to my second supervisor, Dr Rowena Kasprowicz, who reinforced my research design through her insightful suggestions, feedback and sustained support. I could not have withstood the struggles of academic life and research without either of my supervisors. I also take this opportunity to acknowledge the vital contribution made by Dr Louise Courtney and Professor Jeanine Treffers-Daller during the early stages of my study. Similarly, I would like to thank Dr Daisy Powell, Dr Holly Joseph and Dr Anthony Zhang for running the SPSS course. They provided me with invaluable guidance and suggestions in my first phase of analysing the quantitative data.

In addition, it gives me great pleasure to thank the volunteer teachers and students in the English Language Center (ELC), who participated in this study, freely giving of their time and co-operating during the data collection. In this regard, I would equally like to thank the College Dean, the Head of ELC, and the Level One Co-ordinator, for facilitating access to the classrooms. Special thanks are also due to the IT technicians and the ELC's Self Access Centre (SAC) staff for their technical assistance and efforts in enabling the intervention to run smoothly; these were greatly appreciated.

Aside from the above, my heartfelt gratitude goes out to friends and associates who made themselves available to support me during my studies, such as my close friend, Maali Aloudah, who never ceased to offer her encouragement, moral support and friendship during these years. Special thanks are also due to Mr. Musallam Al Amri, the Oman Cultural Attaché, who never failed to respond to my inquiries, both before and after I travelled to the UK. The valuable assistance provided by Ms Maryam Al Khusaibi, Assistant Director General of Scholarships at the MOHE, was greatly appreciated. I would equally like to thank Mr Cass, my former tutor and then boss at Dhofar University, during the time I worked there; he continually encouraged me at all stages of preparing my thesis.

Finally, I am profoundly indebted to all my relatives and friends in Oman and the UK for listening when I needed to talk, encouraging me when the going got tough, and giving helpful advice when I needed direction. Most importantly, there are no words to express my heartfelt gratitude to my wonderful parents, brothers and sisters, who have never stopped believing in me, and by no means least, to my dearest husband, Mohammed Al Dhahab and our children, Fatima, Abdullah, Sara and Tamim, for their understanding and patience throughout my study journey.

ABSTRACT

This study aimed to investigate the impact of time distribution (i.e. spacing between practice sessions; intensive versus spaced) on vocabulary learning and retention in low-ability learners of L2 English using a Computer-assisted Language Learning tool (Quizlet), applied in actual practice at an Omani College of Technology. It also explored teachers' and students' perceptions of the use and implementation of Quizlet, in terms of its perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI). A quasi-experimental design was adopted. The sample was drawn from six intact Level 1 classes from the English Language Center (ELC) at the College, comprising four teachers and 96 students. Students were allocated to three groups: intensive (1 day between practice sessions); spaced (7 days between practice sessions); and a test-only control group. The Experimental Groups (Intensive and Spaced) received four practice sessions (of 20 minutes each) to rehearse the target words. Baseline tests (the 2,000-5,000 vocabulary level test, a background questionnaire and working memory test) and three performance tests (pre-, immediate post- and delayed post-tests) were completed by all three groups. To investigate the participants' perceptions of using Quizlet, interviews were conducted with the teachers, and the questionnaire was administered to students from both Experimental Groups. Both Experimental Groups demonstrated significant improvement in the immediate post-test, which they maintained in the delayed post-tests. Both groups scored significantly higher than the Control Group, who did not demonstrate any change over the three performance tests. Moreover, no significant differences were found between the Experimental Groups. Therefore, the findings suggest that there is no optimal spacing for better retention. In addition, both the teachers and students generally considered Quizlet to be easy to use and beneficial for vocabulary learning and expressed their willingness to use it in future.

DECLARATION OF ORIGINAL AUTHORSHIP

I confirm that this is my own work and the use of all materials from other sources has been properly and fully acknowledged.

Signature:

Print name: Muna Hussain Muqaibal

Date: 03/01/2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AWL	Academic Word List
BI	Behavioural Intention
CALL	Computer-assisted Language Learning
CATSS	Computer Adaptive Test of Size and Strength
CAVL	Computer-assisted Vocabulary Learning
CGE	Computer-generated Exam
DV	Dependent Variable
EAP	English for Academic Purposes
EFL	English as a Foreign Language
ELC	English Language Center
ELCS	English Language Curriculum Section
ET	Educational Technology
ETC	Educational Technology Center
FP	Foundation Programme
GSL	General Service List
HEIs	Higher Education Institutes
ICT	Information Communication Technology
IG	Intensive Group
IGT	Intensive Group Teacher
ILH	Involvement Load Hypothesis
ISI	Intersession Interval
IT	Information Technology
IV	Independent Variable
KG	Kindergarten
L1	First Language
L2	Second Language
LTM	Long-term Memory
PC	Personal Computer
PEOU	Perceived Ease of Use
PU	Perceived Usefulness
RI	Retrieval Interval
RQ	Research Question
SG	Spaced Group
SGT	Spaced Group Teacher
SLA	Second Language Acquisition
SQU	Sultan Qaboos University
SSPS	Statistical Package for the Social Sciences
STM	Short-term Memory
TAM	Technology Acceptance Model
TFA	Technique Feature Analysis
TPACK	Technology, Pedagogy, and Content Knowledge
UK	United Kingdom
US	United States
VLT	Vocabulary Language Test
VST	Vocabulary Size Test
WM	Working Memory
WMT	Working Memory Test

CHAPTER ONE: INTRODUCTION

1.1 Introduction

Chapter One investigates several themes to identify the research problem. In particular, it addresses the importance of learning English in Oman and the progress of teaching English as a foreign language (EFL) and use of technology in that setting. This chapter also outlines the challenges of English language learning and spaced practice, especially in the relevant context, while also presenting the problem statement. In addition, the research aims and the research questions are presented in this chapter, with an overview of the current thesis being provided in its final sections.

1.2 Identifying the Problem

Mahu (2012) highlights the fact that English is a mandatory school subject in many countries and is now the most commonly spoken language worldwide, as it can be understood by one fifth of the world's population. The above author also notes that there are nearly 800 million English speakers globally, with around 400 million native speakers, 300 million non-native speakers whose English is the second language (L2) and 100 million who speak English as a foreign language (EFL) (Mahu, 2012). One of the main reasons for this global spread of the English language is technology.

To elaborate on this point further, English has become the main language medium for information communication technology (ICT), accompanying the massive revolution that has taken place in this area, namely through the internet and ubiquitous use of computers (Al-Jadidi, 2009). Referring to Hasman's (2004) statistical data, it is estimated that more than 85% of emails sent around the world are printed in English and 90% of the world's retrievable information is stored electronically in English. These figures will have surely increased significantly since the time of the above study, but serve as an indication of the vast reach of the English language, which also includes various fields of knowledge and science. For example, Hasman (2004) states that within a single decade, more than 70% of the world's scientists have come to read in English, with the English language now being utilised by numerous researchers and scholars as the language of the sciences. Drubin and Kellogg

(2012, p.1,399) point to the fact that “English is now used almost exclusively as the language of science”, which demonstrates the extent of English language usage in the field of science.

In today’s world, the English language is even widely used in countries where it is not the official language, because of the requirement to use English in many professional and academic areas (Mahu, 2012). Al-Jadidi (2009) adds that English is considered as an essential feature of global economic growth for every country in the world. Therefore, it has become a necessity for many to learn English, given the acceleration of growth and development in areas such as technology and science, as well as throughout the world’s economies. This glance at the expanding use of English language globally leads us to specifically focus on the point that the terminology and vocabulary used in these areas, i.e. sciences, technology and economy, are frequently used English words. Therefore, the content words in English for Academic Purposes (EAP) have become essential for students to learn to enhance their English language proficiency in their academic settings. Thus, EAP is designed to meet students’ needs in each different programme to enable them to improve their English proficiency in the contexts that they are going to deal with later in their lives (Sharndama et al., 2014) and a key part of the curriculum of the university in this research study.

1.2.1 The Importance of English in Oman

English language has become an essential foreign language in Oman, due to the rising demand for English language skills across many different sectors and fields. Thus, the Oman government acknowledges the official use of English in many organisations in the country’s public and private sectors, such as banks, businesses, the tourist industry and all public and private Higher Education Institutes (HEIs) (Al-Issa, 2006; Al-Jardani, 2012).

In addition, rapid economic growth in Oman has given rise to the need for more English speakers, corresponding to the widespread use of English globally, which has increased over the past few decades. English has, therefore, been adopted as an official foreign language in Oman, whereby the Oman government considers it to be crucial to certain professions, especially in the tourism and business sectors (Al-Issa, 2006). However, its use is especially widespread in the private sector and the petroleum industry, which are connected to the broader sphere of international organisations and companies (Al Riyami, 2016).

Given the government's plans to adopt a policy of 'Omanisation', representing the "systematic and gradual replacement of foreign skilled labour" by Omani citizens (Al-Issa, 2006, p.199), Omani workers are required to have the same level of qualification as foreign experts and skilled workers, in order to be able to fulfil the responsibilities associated with their positions. One criterion consists of English language skills, which are necessary for communication among multi-national staff and to facilitate work with international organisations.

This sense of the importance of learning English has even been reinforced by parents in Oman, who realise that English is "a tool or gate to the future" for their children (Al-Jardani, 2012, p.41). Therefore, many Omani families take advantage of the summer holidays to send their children to language institutes in the country, so that they can improve their English skills. In the case of high-income families, children may be sent outside Oman to places where English is the native language, such as the UK or US, or to countries where it is an L2, such as Malaysia or India.

Therefore, policy makers in Oman's Ministry of Education have paid increasing attention to English teaching methods in the classroom (Al-Issa, 2006; Al-Jadidi, 2009; Ismail, 2011), with the purpose of improving learners' ability to communicate and practise the language in authentic situations, such as when filling out forms or sending emails in English (Al-Jardani, 2012). One key aspect necessary to achieve this goal is the focus on increasing the vocabulary knowledge of learners as an essential tool that enables learners to communicate well in the L2 language. Schmitt (2000, p. 55) asserts that "lexical knowledge is central to communicative competence and to the acquisition of a second language". Therefore, the focus of the thesis is on vocabulary learning, as it constitutes an essential part of the process of language learning as well as teaching (see subsection 1.3.1). What is less clear are the best ways to increase vocabulary knowledge, as there are different theories of vocabulary learning (see subsection 2.3.1). This study, however, aims to explore the activation-based model of the spacing effect (Pavlik & Anderson, 2005) on vocabulary learning and retention using a different learning setting, namely Quizlet, i.e. a flashcard software programme (see subsections 2.4.2.4 & 3.4.4.8). The spacing effect refers to "the effect of time distribution on learning" (Serrano & Huang, 2018, p. 3). It is concerned with an enhancement in learning that can occur when the items to be learnt are distributed in time

(Serrano & Huang, 2018). As the study took place in a college in Oman (see subsection 3.4.1), the following subsections provide a brief background of the development of English language in general, the challenges of English language for Omani students as native Arabic speakers, and vocabulary learning and teaching in particular, in Oman.

1.2.2 The Development of English Teaching and Use of Technology in Oman

Education as it stands today in Oman could be considered as a recent system. Before the renaissance of Oman in 1970, there were only three schools in the country, with just 900 students in total, all of whom were male (Sergon, 2011). After 1970, however, the government made huge strides forward in response to global development in the field of education. Nevertheless, EFL was not introduced as a school subject until 1977, when it was established for Grade 4 (nine-year-olds) and upwards by the Curriculum Department of Oman's Ministry of Education (Al-Jadidi, 2009). With the ongoing development of Oman's education system, English language instruction now takes place from Grade 1 (age six) onwards, with the implementation of a new 'Basic Education' system in 1998 (Sergon, 2011).

On introducing this new Basic Education system, Oman's Ministry of Education has supplied schools with teaching aids and established learning resource centres at all schools nationwide, where computers, visual aids and books can be accessed. However, there tend to be few English language materials for English teachers in these centres, which means that there is a lack of encouragement in schools to benefit from the learning resource centres on a regular basis (Al-Jardani, 2012). The most recent English language curriculum applied in all government schools is based on an in-house course book entitled 'English for Me', used for Grades 1-10 (Al-Senaidi & Wyatt, 2014; Sergon, 2011), and another course book for Grades 11 and 12, entitled 'Engage with English' (Al-Jardani, 2012).

This curriculum was introduced in 1998, resulting from a series of changes that took place in the existing English curriculum (Al-Issa & Al-Bulushi, 2012). This began with two commercially produced courses and ended with the development of an in-house curriculum, which has since been modified five times (Al-Jardani, 2012). The latest changes made were based on the general objectives of Oman's Basic Education system in relation to English language teaching and learning. They are listed as follows: "Acquire knowledge and skills in all areas of the curriculum, including skills in questioning, investigating, critical thinking,

problem solving, and decision making." They also include applying "the skills learned to further studies, work, leisure, daily living and a lifetime of learning" and emphasise the use of "a variety of technologies, demonstrat[ing] an understanding of technological applications, [and] apply[ing] appropriate technologies for solving problems related to [daily life]". Finally, the broad objectives of Basic Education support the use of English "in other subject areas; and present English as an international language and as a means of communication" (ELCS, 2010, cited in Al-Jardani, 2012, pp. 42-43).

Of some concern, however, Al-Mahrooqi et al. (2016) observe that the current English textbooks used for Grades 1-12 fail to cover the essential word lists, namely, the General Service List (GSL) and the Academic Word List (AWL), which were proposed by Nation (1990, 2001) to provide a text coverage ratio of 95% and 98%. Moreover, their analysis of these English textbooks using WordSmith Tools (Scott, 2008) and RANGE (Heatley et al., 2002), two kinds of concordance software, revealed that a large proportion of their vocabulary is not contained in these fundamental lists. They concluded that there exist irregularities in the load of vocabulary in the current textbooks, with vocabulary being presented at inappropriate levels. In the view of many teachers, weaknesses in the textbooks hinder the effective teaching and learning process (Al-Mahrooqi et al., 2016). These authors highlight the importance of technology through utilising concordance software to evaluate vocabulary in the textbooks. This kind of software programme helps to show how vocabulary is consistently distributed in the English textbooks of all the grades and to check the presence of the essential word lists, such as GSL and AWL, in these textbooks. These programmes can help the curriculum designers to monitor and improve the English curriculum in terms of vocabulary knowledge through including the words needed in each grade without overwhelming students with unnecessary words. This evaluation helps teachers to recognise the target vocabulary that match learners' needs and current levels in order to facilitate the process of English teaching and learning for both teachers and learners. The lack of essential vocabulary in English school textbooks may also point to one of the core reasons for the limited vocabulary size of many Omani students at university level, despite having received many hours of English instruction at school. Therefore, it is important to consider whether alternative or supplementary methods of instruction could be beneficial to provide instruction in the necessary vocabulary at an appropriate level.

In Omani state schools, students have five to seven English classes per week, each for a duration of 40 minutes, meaning that they receive between three and four hours of instruction per week. In the Basic Education system, students have both English and Arabic lessons in school, starting at Grade 1 (age six) and progressing to Grade 12 (age 17-18) (Al-Jardani, 2012). Therefore, Arabic is the default language of learning and instruction in state schools, in contrast to Oman's HEIs, where English is the medium of teaching. Despite this level of input, many students are just not finding learning English easy and need alternative types of support.

Moreover, most public sector institutions under Oman's Ministry of Higher Education utilise English as "the medium of instruction [especially] in all the science-based majors" (Al-Issa, 2006, p.201) – as is the case in "the Institute of Health Sciences, Higher Colleges of Technology, the College of Banking and Financial Studies, the College of Sharia and Law, the Colleges of Education, Sultan Qaboos University and the Royal Air Force of Oman Academy" (Al-Issa, 2006, p.200). However, English is not only used in public sector institutions, but also in private schools and HEIs. In private schools, for example, English is taught from an early stage, starting with Year One of kindergarten (KG1), for three or four year olds. In addition, private colleges and universities utilise English as the language of teaching at all stages of higher education in both arts- and science-based subjects (Al-Issa, 2006).

Including both the private and public sectors, HEIs comprise over 20 colleges and universities in Oman. In this country, Educational Technology (ET) and e-learning are gradually being integrated into the educational system. There is much encouragement by academic staff and administrators in HEIs toward incorporating ET in the process of learning and teaching (Al Musawi, 2007). Al Musawi (2007) claims there are two main reasons for the dynamic development in ET in terms of staffing, funding and equipment at these institutions: "improvements and new acquisition of modern technology and software and improved administration support for media use in teaching" (p. 396-397). Surfing the internet is now available for students so they can easily get information and knowledge and connect them with their counterparts in any place on the earth (Al Rawahy, 2001). This positive view of using technology in the process of learning and teaching at HEIs encouraged the researcher to integrate technology through using a software programme (Quizlet) (see subsections 2.4.2.4

& 3.4.4.8) to support vocabulary learning at college. Generally, vocabulary software programmes provide learners with an audio-visual exposition which supports the mental association between the forms and meanings of the words (Nation & Webb, 2011). This research was, then, carried out in the English Language Center (ELC) in a Technical College, which is a public HEI in Oman (see subsection 3.4.1).

The College aims to provide high quality student services through Information Technology support. The Educational Technology Center (ETC) is responsible for delivering and running IT facilities and resources at the College in order to support learning in technological education. These IT facilities and resources include a network operation centre, library facilities, e-learning laboratory, multimedia facilities in the Multipurpose Hall and photocopying centre. The College has a total of 28 air-conditioned computer laboratories connected to the internet for innovative classroom instruction. They are also supplied with up-to-date models of computers, high-end equipment and multimedia projectors. Some laboratories are provided with smartboard technology (SCT, 2018). In the ELC, all level teachers are required to use Moodle, an online course forum. Teachers enrol their students into Moodle to utilise the online materials, supplemented by the level coordinators and teachers. The categories in Moodle include all the courses of the college departments and centres, and also other online services. The ELC category involves the four ELC levels. For example, Level 1 includes short stories, grammar, Pathways vocabulary log, timetable and Level 1 delivery plans. The Level 1 link also provides students with all four English language skills (Reading, Listening, Writing and Speaking). Each skill course on this link is divided into course book units, and each unit includes several activities, links or videos to promote students' learning. Currently, all Reading and Listening classes in the ELC are held in the laboratories. This has been incorporated since 2018 after implementing the new English textbooks, namely "Pathways" (Davis, 2017). Teachers have been asked to create their own accounts on the "MyELT" website to benefit from online supplementary activities and materials along with the new curriculum Pathways. Although most of the teachers have not used this website with their students yet, they had a workshop about how to use it in their classes. This could be the next phase for the ELC to integrate more educational technology in all the courses at the ELC in the next academic years. With all these facilities and plans, the College is an ideal environment supporting educational technology in the process of learning

and teaching. However, it is important to know how to make the best use of technology. This thesis aims to explore the use of these facilities available at the College to support the ET effectively in learning and teaching English vocabulary in the ELC's Foundation Programme.

Almost all post-Foundation subjects and specialisations are taught in English, with English consequently being the language of instruction at most Omani universities and colleges, and students being expected to write their assignments in English. Consequently, the English language is an essential component of the learning process in Oman. However, Omani students, as native Arabic speakers, face many challenges in their English language learning and vocabulary acquisition, which are explored in the next subsection.

1.2.3 The Challenges of English Language and Vocabulary Learning for Omani Students as Native Arabic Speakers

Arabic is the official language of most Middle-Eastern countries (Abdelgadir & Ramana, 2016), including Oman. There are numerous challenges facing native Arabic speakers (Arabic L1) when attempting to learn English. The first of these arises from the fact that Arabic and English are from two different language families, with Arabic being Semitic and consisting of 28 consonants and just three written vowels, namely the 'long vowels'. The other vowels are not written, but may be indicated using diacritical marks (Nasr, 1963, cited in Saigh & Schmitt, 2012). Therefore, Arabic is more consonantal (Cook & Bassetti, 2005), with "almost a 1:1 phoneme-grapheme representation" (Saigh & Schmitt, 2012, p.26).

In contrast, English is an Indo-European language with 24 consonant sounds and 20 long and short vowel sounds. In particular, these short vowel sounds are not considered important in Arabic, and are ignored in many cases by Arabic speakers when they write or speak (Abdelgadir & Ramana, 2016). Meanwhile, English applies an alphabetic writing system (Cook & Bassetti, 2005), with consistent and inconsistent phoneme-graphemes, or more complex representations. Thus, English is relatively more opaque than Arabic (Saigh & Schmitt, 2012). In addition, Arabic learners of English are likely to encounter other problems that are not the focus of this study, such as reading cursive writing, although Arabic speakers can also read from left to right. Moreover, they often ignore the use of punctuation pointers and fail to differentiate between lower and upper cases (Abdelgadir & Ramana, 2016).

English, therefore, presents many challenges for native Arabic speakers, which some learners find difficult to overcome.

Consequently, Omani students as L1 Arabic learners face much difficulty in vocabulary processing and spelling, compared with other ESL learners with different L1 backgrounds. This is due to the impact of the conventional spelling system and literacy skills, i.e. the English word recognition and the orthographic knowledge of English words (Fender, 2008; Milton, 2009). As this study mainly focuses on vocabulary learning, the most important difference between Arabic and English concerns their respective systems for writing word forms, in that unlike Arabic, English is not considered a transparent language (Saigh & Schmitt, 2012). This can make it confusing for Arabic-speaking English learners, especially with regard to the vowels in English, which often go unnoticed (Alsadoon, 2015).

In addition to Arabic language, most students at the College speak an unwritten local language or dialect (see subsection, 3.4.2). These different forms of speech affect students' pronunciation. This issue has received limited attention in the research literature and has been discussed very recently by one author, Attamimi (2018), who points out that students who speak Jabbali, which is spoken by most of the students at College, encounter problems with distinguishing between the soft /g/ sounds and the hard /j/ sounds, "so that "girl" becomes /'jɜ:l/, or "language" becomes /'læŋwɪdʒ/" (p. 8). These groups of students also have problems with the /v/ sound, which does not exist in Arabic and Jabbali. Students often avoid pronouncing this letter or replace it with the /f/ sound. Such problems can affect students' spelling, pronunciation, and recognition of English written and aural word forms. There is a lack of research about the specific challenges faced by Omani students who speak various dialects and come from multicultural groups. All these issues are particularly pertinent to students learning English, and especially to the process of learning vocabulary, because of the influence of their L1, which can be so different from the L2 being learned. This difference is referred to as 'linguistic distance' (Chiswick & Miller, 2005).

There are several methods of learning vocabulary. In the ELC at College, students are provided with vocabulary lists, called Vocabulary Log (see Appendix 14). Words in Vocabulary Log are elicited from their course books to facilitate their understanding of the subject matter when they encounter these vocabulary items during their lessons. Students are

asked to find out the L1 equivalents of all words provided in the Vocabulary Log, identify their parts of speech and put them in sentences. The responsibility of the teacher is to check that students complete this vocabulary assignment. This assignment should be done by the students at home; however, at best its value is mixed, as it is often not clear that all the work is done by the student amidst the widespread belief that a minority of students copy the answers from each other. So, it is difficult to guarantee that students fully benefit from this assignment. Teachers need more reliable assistance to assess whether the students are actually doing their own work. Therefore, this thesis aims to generate a greater understanding of vocabulary learning and retention in real practice at College through adopting an explicit method of vocabulary teaching using two different spaced practices and utilizing technology available at the College.

1.2.4 Spaced Practice and Lag Effect

A large body of research, in the field of cognitive psychology, has been conducted to explore the impact of time distribution (spaced versus massed) for practice/study sessions on learning (Suzuki & DeKeyser, 2017). It has been shown that there is generally an advantage for distributing study sessions in spaced sequences, rather than massing study into a single session (e.g Carpenter et al., 2012; Cepeda et al., 2006). In massed practice, all the learning activities are completed in one session, while in spaced practice the activities are spread out over time, in multiple study sessions with a gap in between each session (e.g. Rogers, 2017; Rohrer, 2015; Serrano, 2012). Serrano and Huang (2018) have taken note that the majority of testified results on the spacing effect have relied on laboratory studies, where students received their treatment via computer screens, usually to learn vocabulary at different time retrievals, frequently in one study session in cognitive psychology experiments.

On the other hand, in the related literature on Second Language Acquisition (SLA), the investigations have distributed learning over multiple study sessions with different lags between these sessions (Rogers, 2017). It is also important to note that in typical educational settings, learning tends to be distributed over several classes (Rohrer, 2015). Therefore, rather than comparing massed versus spaced study (i.e. the spacing effect), what is of interest in the SLA is whether the amount of time *between* study sessions has an impact on learning and retention (Rogers, 2017).

The comparison between two different spacing schedules (i.e. longer versus shorter) has revealed what is called the "lag effect" (Rogers, 2017), which is a particular focus of this study. The lag effect refers to the finding that the recall of words to be learned is better when there is a longer interval between the learning sessions of those words. There are a small number of other studies (Bird, 2010; Kasprowicz et al., 2019; Rogers, 2015; Suzuki & DeKeyser, 2015; Suzuki, 2017) which have looked at longer and shorter spacing for L2 grammar learning. These studies have found mixed results. Some have found that longer spacing yielded better retention (e.g. Bird, 210; Rogers, 2015), whereas other studies have found that shorter spacing had a significant advantage (e.g. Suzuki, 2017; Rogers & Cheung, 2018), and others have got no statistically significant differences between shorter and longer lags (Kasprowicz et al., 2019; Küpper-Tetzel et al., 2014). It should be mentioned that to the current researcher's knowledge, Kasprowicz and her colleagues' study is the only lag effect study that had a control group. Thus, this thesis aims to explore the effect of time distribution using two different time sequences and including a control group for each time sequence (see section 3.3.4), which is one of the original contributions to the knowledge of this research.

Only a small number of studies and meta-analyses have given attention to the effectiveness of time distribution between sessions on enhancing learning (e.g. Cepeda et al., 2006; Rohrer & Pashler, 2007). In regard to these studies, the intersession interval (ISI) refers to the amount of time between study sessions, and the retention interval (RI) indicates the gap between the last study session and the test. Rohrer and Pashler (2007) studied the interaction between the ISI and the RI and concluded that the length of the ISI (i.e. time between the sessions) is determined based on when the testing time takes place. Rohrer and Pashler suggest that the ratio of the ISI should be between 10% and 30% of the RI. Therefore, it is important to look at what this means in practice in a real classroom to explore the impact of the lag effect using short and long ISIs on vocabulary acquisition and retention.

1.3 Context of the Problem

Based on the enhanced education system in Oman, students undertaking their Basic Education (Grades 1-10) are now expected to recognise 4,500 words, and when they graduate from the Post-Basic Education stage (Grades 11-12), it is anticipated that they will be able to recognise 6,000-7,000 words in English (Curriculum Framework 10, cited in Sergon, 2011). According

to Nation (2013), learners should be able to cover 95% of an authentic reading text, if they recognise 3,000-4,000 word families. If they know 6,000-9,000 word families, their understanding of authentic reading texts rises to 97%.

With the highly optimistic expectation of word gain by the designers of the Curriculum Framework at the Omani Ministry of Education, it is anticipated that by Grade 12 (age 17-18), students can read and speak English well, as the designers expected students, after a long time of learning English at school, to have an adequate level of vocabulary knowledge. However, this expectation of Omani learners does not match the reality. Laufer (2000) states that the vocabulary size of Omani EFL students at university level, after having received around 1,350 hours of English instruction, amounts to just 2,000 words. Similarly, the extent of vocabulary knowledge for Japanese EFL students in university, for example, is between 2,000 and 2,300 words, after being exposed to between 800 and 1,200 hours of English instruction. Meanwhile, Indonesian students at this stage of their education, following 900 hours of English tutoring, can recognise only 1,220 words (Laufer, 2000).

In the context of Oman, although pupils in government schools begin studying English as early as Year 1, Al Jabri (2008) noted from his observations and interviews with English teachers in one of Oman's public sector schools, that students had observable difficulties with vocabulary, which hindered them from speaking, listening, reading and writing effectively. Al Jabri found that Omani students at government schools have poor recognition of vocabulary items. They have the problem of grasping new words from the context or understanding simple sentences that include high frequency vocabulary items. It was also noticeable that many students did not follow the teachers in their EFL classes. The teachers in Al Jabri's study found that the vocabulary in the English textbooks in the students' current grades was much higher than their actual level of vocabulary proficiency. They attributed the students' low vocabulary proficiency to the lack of motivation towards learning English language, the length of the English curriculum at school and the lack of provision of strategies of learning vocabulary outside the classroom. However, Al Jabri considers that many English teachers in Oman do not give much attention to the explicit teaching of vocabulary during their classes, as they believe that vocabulary can be acquired automatically throughout several class activities given to them. Al Jabri (2008, p.1) states that "many English language teachers in Omani government schools do not provide their students with

the main tool for communicating in a language, which is vocabulary". Although Al Jabri's study is over a decade old, the teachers' opinions about the high proficiency level of vocabulary provided to the students are consistent with the recent findings of Al-Mahrooqi et al.'s (2016) study about the vocabulary in the school textbooks (see subsection 1.2.2). Thus, the priority that teachers place on vocabulary learning can also have influence on the low level of English proficiency among Omani students, as there is a strong relation between vocabulary knowledge and English language skills (see subsection 1.3.1). Several studies indicate that having sufficient vocabulary knowledge reflects good performance in the four language skills (e.g. Nunan, 1999; Smith, 2003).

New students in all HEIs are required to take a placement test, so that they can be allocated to the right level on the Foundation Programme. This placement test measures students' abilities in English, Mathematics and IT. For this purpose, HEIs administer an in-house placement test, designed according to their learning outcomes as HEIs (Al-Mamari, 2012). However, only very few students tend to pass these placement tests and enrol directly on a college or university programme. Al Mahrooqi (2012, cited in Kamanpoori, 2014) refers to a study that she conducted for His Majesty the Sultan's Trust Fund for Strategic Research in May 2012, indicating that 80% of students undertake English courses in the Foundation Programme, when accessing higher education. This figure is based on a study conducted on a sample of 8,000 students, producing findings that indicate a weakness in students' English language skills (Kamanpoori, 2014).

HEIs are required to source English language programmes and develop English instruction to help students improve their English skills, as this is the language of instruction at most colleges and universities in Oman. Therefore, a lot of consideration needs to be dedicated to improving students' vocabulary knowledge, especially in terms of spacing effects in vocabulary learning using technology, which this study aims to investigate.

1.3.1 Statement of the Problem

Wallace (1982, p.9) argues that vocabulary plays a central function in learning a language in general, "because there is a sense in which learning a foreign language is basically a matter of learning the vocabulary of that language". Thus, vocabulary is considered to be one of the most important aspects of language learning as a second or foreign language. It plays an

important role in all four language skills, i.e. writing, speaking, listening, and reading (Nation, 2013). Without an adequate level of vocabulary knowledge, successful communication can be very difficult to achieve. Vocabulary is a dominant element of communicative competence and language acquisition (Schmitt, 2000), for the fundamental reason that conceptual knowledge and thoughts need to be converted into spoken or written lexical items in order to be expressed to others. Therefore, vocabulary forms an essential part of human communication. It is also important for mastering writing skills, in which words embody the orthographic form of ideas. Listening comprehension also relies on vocabulary knowledge; when learners have a deficient level of vocabulary knowledge, they will struggle to comprehend the spoken language (Droop & Verhoeven, 2003). Further, studies have shown a positive correlation between vocabulary knowledge and reading comprehension (e.g. Biemiller & Boote, 2006; Staehr, 2008), in which the direct and explicit teaching of vocabulary both enhances learners' vocabulary knowledge and improves their levels of reading comprehension (Hunt & Feng, 2016). Therefore, Hunt and Feng (2016) emphasize the importance of using effective methods specifically in teaching vocabulary to engage learners in the process of vocabulary learning through explicit vocabulary teaching, such as previewing and reviewing the reading text, reading aloud and using multimedia.

However, vocabulary is not yet explicitly taught to any great extent in language classes at the College in Oman. Instead, teachers use the traditional approach of providing students with lists of vocabulary drawn from their courses of study. These are given to students at the beginning of the semester, so that they can study them at home. Tests are then administered on selected items from the lists at a later stage. In addition, Mehring (2005) supports the notion that vocabulary learning requires extensive practice and time, as it is considered as a continual learning process. The extent of this is quite varied for individual students. The lack of practice and absence of explicit vocabulary teaching in the classroom at the College may be a possible reason for the low level of word knowledge amongst Omani English language learners. There is little evidence that this method is working for a large minority of students, and therefore, alternative methods need exploring.

Transferring from traditional to digital methods of teaching vocabulary has become a necessity, as we live in a digital age and are dealing with a digital generation. However, the use of technology will only make a positive difference to learning if it can be established how

to use it effectively (Higgins et al., 2012). Referring to a meta-analysis of 48 research studies conducted over a period of 40 years, Higgins and colleagues (2012) concluded that there was a consistently positive impact of using technology to support learning. Therefore, teachers need to involve technology in their vocabulary teaching, so that will appeal to the new generation of learners. Computer-assisted Language Learning (CALL) and Computer-assisted Vocabulary Learning (CAVL) programmes are consequently considered to be good environments for learning vocabulary, where students can learn via multiple media, such as sound, images and video (Lin, 2010). For example, Quizlet is one such well-known online flashcard programme (http://quizlet.com/), which is attractive for colleges and universities because it is available free of charge. This is explained in detail in Chapter Three of this thesis. However, not just the integration of technology in learning vocabulary is important, but also how to retrieve and memorize word meanings. A large number of cognitive psychological studies have supported using spaced presentation (see subsection 2.3.3) across several learning sessions to achieve better retention of word meanings (Cepeda, et al., 2006; Goossens et al., 2012).

Nation (2013) argues that flexibility in presenting words in a different order, a practicable size, and a spaced repetition technique can render flashcards more effective for memorising new words, as well as facilitating their use. Therefore, for the current study, an opportunity sample of students with low levels of English proficiency was selected from a college in Oman. This study seeks to provide evidence of the effectiveness of using technology (i.e. Quizlet) and vocabulary teaching in the classroom comparing between different spacing practice (intensive versus spaced), in order to improve learners' retention of new vocabulary in immediate and delayed post-tests.

1.4 Research Aims

This study aims to compare the effectiveness of time distribution using two different time sequences (intensive versus spaced) in promoting vocabulary learning and retention in low ability learners, and to explore the teachers' and students' perceptions on the use and implementation of Quizlet.

1.5 Research Questions

1. Is Quizlet (a Computer-assisted Language Learning [CALL] tool) an effective programme for promoting vocabulary learning amongst low ability learners in the classroom?

2. Does the time distribution of the practice sessions (intensive versus spaced) moderate the benefits of using Quizlet to promote vocabulary learning and retention amongst low ability learners?

3. What are the perceptions of teachers and students regarding the use and implementation of Quizlet?

1.6 Overview of the Thesis

This thesis contains seven chapters. The first chapter, Chapter One, has already identified the research problem, including the importance of learning English, the development of teaching English as a foreign language (EFL) in Oman and integrating technology in Omani educational institutes. In addition, the challenges of English language learning and spaced practice were outlined. Finally, the chapter presented the problem statement, the research aims, and the research questions.

The second chapter, Chapter Two, reviews three main areas of the research literature. The first section presents an overview of vocabulary knowledge. It includes knowing a word, receptive and productive knowledge, defining and counting words, and vocabulary size. This second section covers the topic of vocabulary learning and the main underpinning theory applied in the present study, namely explicit vocabulary learning theory. This section also briefly presents theories of memory and L2 vocabulary learning, including levels of processing, the memory system, the cognitive perspective of vocabulary learning and technique feature analysis (TFA). The third section then addresses vocabulary teaching and practice, focusing on distribution of practice and the impact of lag effects (i.e., longer versus shorter spacing) on learning, and the use of technology for vocabulary learning, beginning with a definition of CALL and then giving an overview of CALL instruction and vocabulary learning, as well as an overview of the literature on Quizlet. With this aim in view, the chapter presents a review of related studies.

Chapter Three explains the research methodology implemented in this study. First, it identifies the research questions and hypotheses, rationale for the selected research methodology, and the research methodology. The data analysis then describes the marking scheme for the performance tests (pre-, immediate post- and delayed post-tests), the quantitative data analysis, checking assumptions and qualitative data analysis. The chapter also outlines the reliability and validity, and ethical issues.

Chapters Four and Five then present and review the findings of the data analysis. The overall aim of this research is to ascertain the impact of different spaced practice (longer versus shorter spacing) using Quizlet to facilitate vocabulary learning and retention for low level learners, and to investigate participants' perceptions of using Quizlet in vocabulary learning and teaching, in terms of perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI).

Chapter Six discusses the findings of each research question and links them to the theories and other findings of the relevant studies in the chapter of the literature review. The anticipations and possible reasons are also presented in this chapter to interpret the differences in findings between the present study and previous empirical studies.

The final chapter, Chapter Seven, summarises the presents study and its main findings, drawing conclusions from the research results, presenting the limitations and making recommendations for further research. Finally, the chapter discusses the pedagogical implication of the study in the classroom, and identifies the contributions of this study.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

The purpose of this chapter is to provide an overview of three main streams of literature related to this research: an overview of vocabulary knowledge, vocabulary learning, and vocabulary teaching and practice. The section dedicated to vocabulary knowledge briefly covers the fact of knowing a word, productive and receptive knowledge, defining and counting words, and vocabulary size. The following section on vocabulary learning includes theories of vocabulary learning, memory, levels of processing and Technique Feature Analysis (TFA). Finally, the third section on vocabulary teaching and practice presents an outline of distribution of practice, involving lag effects, i.e. the impact of longer versus shorter spacing on learning, and Computer-assisted Language Learning (CALL) for vocabulary learning, including an overview of the literature on Quizlet and the Technology Acceptance Model (TAM).

In these above-mentioned sections, the literature relating to specific key theories and concepts will be reviewed. The chapter then concludes by outlining the direction planned for the current study.

2.2 Vocabulary Knowledge

Vocabulary knowledge is central to this thesis, because of its vital role in fortifying language proficiency and enhancing communication skills (Masrai & Milton, 2016; Milton, 2009; Nation 2013). According to Nation (2013), it involves many varieties of language aspects, including knowing a word, identifying ways of counting vocabulary, and measuring vocabulary size. Each of these processes will now be discussed in more detail.

2.2.1 Knowing a Word

According to Nation (2013), the fact of ‘knowing’ a word is classified into ‘receptive’ and ‘productive’ knowledge. Receptive knowledge consists of comprehending language input from others, i.e. through listening and reading. Conversely, productive knowledge involves the production of language in either written or spoken form, in order to deliver the meaning to others. Receptive and productive vocabulary knowledge cover different aspects of knowing a

word, in terms of its meaning, forms and uses, as captured in Nation's treatment of knowing a word in Figure 2.1.

Figure 2.1
"What Is Involved in Knowing a Word?" (Nation, 2013, p.49)

Form	**spoken**	R	What does the word sound like?
		P	How is the word pronounced?
	written	R	What does the word look like?
		P	How is the word written and spelled?
	word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express the meaning?
Meaning	**form and meaning**	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	**grammatical functions**	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use (register, frequency…)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

Note: R = receptive knowledge, P = productive knowledge.

Therefore, Nation's model for the concept of 'knowing a word' is important for this study, as it can be applied as a guideline for the measurement (Milton, 2009) and evaluation (Maskor & Baharudin, 2016) of different aspects of vocabulary knowledge. Two aspects of Nation's model, form and meaning, will be focused on for the purpose of this study, although it is acknowledged that its third aspect, use (including the context in which words are used) is also important to take into account in knowing and learning vocabulary. However, the context of the word used in the sentence is beyond the scope of this study. The two main classifications of lexical knowledge, productive and receptive, will be explained in the following subsection.

2.2.2 Productive and Receptive Knowledge

Many linguistics scholars with a special interest in vocabulary classify different types of vocabulary knowledge, based on the 'productive' (active) skills of writing and speaking, and the 'receptive' (passive) skills of reading and listening (Maskor & Baharudin, 2016; Milton, 2009; Schmitt, 2010). Schmitt (2010) argues that learners are more capable of acquiring receptive knowledge than they are of gaining productive knowledge and points out that there is no clear relationship between these aspects of vocabulary knowledge.

Henriksen (1999) takes a somewhat different stance and also allocates three dimensions of knowing a word: (1) starting with partial knowledge, which involves ascertaining the exact meaning of the vocabulary, (2) moving on to deeper vocabulary knowledge, where various aspects of word knowledge and the relationships between words are recognised, and finally, (3) acknowledging receptive and productive knowledge, where learners demonstrate their knowledge of a word.

Contrary to this, Nation (2013) adopts a broader perspective and lists three aspects and features of each, including meaning (associations, form and meaning, concepts and referents), form (word parts, written and spoken), and use (collocations, constraints of use and grammatical functions), in order to define what is meant by 'knowing' a word, with productive and receptive knowledge being involved in each part (see Figure 2.1). Understanding word features (form, meaning and use), and recognising the relationship between these features and parts, enables teachers to design appropriate tasks and learning strategies for use when teaching vocabulary in the classroom, thus enhancing both productive and receptive aspects of learners' vocabulary knowledge (Maskor & Baharudin, 2016). Therefore, key researchers do not all agree on the specific aspects crucial to learning vocabulary, but they do agree that vocabulary activities in this research involve both productive and receptive knowledge, focusing on meaning and on both the written and oral forms of target words, as part of knowing a word.

2.2.3 Defining and Counting Words

The second aspect of vocabulary knowledge involves defining and counting words. The act of counting words is fundamental to this thesis as a means of identifying learners' language

productivity and comprehension. First, however, it is required to be established what exactly is meant by the term 'word' and by 'counting words'. Words are forked components associated with many different levels and systems in a language, while each word embraces multi-faceted information and degrees of knowledge – with regard to learning its form and meaning, and understanding its usage (Nation, 2013). Drum and Konopak (1987) define a word as "an acoustic configuration of speech and a written rendition (more or less) of these sounds [which] comes or is assigned to refer to things, events, and ideas arbitrarily" (p.73). In simple terms, a word is a sound or written form used to convey the meaning of something that can be seen, touched, felt, thought about or imagined. However, it is rarely that simple.

Drum and Konopak (1987) emphasise that it is culture that determines the meaning of each word. Words have different meanings depending upon the cultural context that the individual is experiencing. Therefore, the unique meaning of a word reflects the understanding of that word within the context of the culture of the individual. For example, the English words *oven* and *bakery* have the same L1 equivalent '*furn'* in the Arabic language (Jiang, 2002). Therefore, identifying the meaning of a word will reflect an understanding of how to communicate knowledge within the framework of the learners' immediate culture. This is relevant for this thesis because it is important to be aware of the participants' culture that the study takes place in. This understanding of their culture helps the teachers and researcher to find out other tools, such as images, to convert the correct meanings of L2 words that could have the same L1 translation, as explained in the above example of '*furn'*. Hence, vocabulary learning goes beyond merely encountering the word and then processing its form; learning a word means that learners are also required to understand its meaning, culture uses and be familiar with its pronunciation and usage in a sentence or context (Courtney, 2014).

Based on the above discussion, words take different forms and can carry a range of meaning and types of meaning. Therefore, it is important for this thesis, as explained above, to investigate how they should be counted. Referring to Milton (2009), the term 'word' gives rise to specialist definitions, based on how words are counted in a sentence or text, such as according to type, tokens, lemmas or word families. Explaining in greater detail, Nation (2013) lists four different techniques for counting words. The first of these involves 'tokens' or 'running words', where every single word in a line, page or book is counted, in relation to a

person's average reading and speaking speed per minute. Even a repeated word is counted twice or more, depending on how many times it is repeated when using tokens as a means of counting words. Where 'type' is considered, however, repeated words are not counted. The aim of the latter approach is to determine the number of words that must be read in a certain book or to count the number of words in a dictionary.

In contrast, lemmas and word families are groups of words that relate to each other, gathered under one word or a single unit. Using these approaches to count words will markedly reduce the number of corpus units. For example, Bauer and Nation (1993) counted 37,617 lemmas out of 61,805 tagged types in the Brown Corpus, recording a decrease of 40% (cited in Nation, 2013). To be more precise, a lemma is a group of related words under a single lexical heading. These related words represent lexical inflections and short forms (for example, *can't*). Meanwhile, a word family contains a main word plus its derivations and inflections. Determining the unit of vocabulary to use for counting will depend on how useful it is to the language-learning purpose (Nation, 2013).

Vocabulary counting is important for this study, in order to be able to estimate vocabulary size. For this research, Nation's fourth technique for counting words, namely the 'word family', has been adopted as the best method of defining vocabulary size. In this way, the researcher will be able to gain a meaningful perspective of vocabulary size for the learners participating in this present study, which will in turn enable the target vocabulary to be determined. More information about the selected target vocabulary will be detailed in subsection 3.4.4.9. However, the next subsection will look at estimating vocabulary size.

2.2.4 Vocabulary Size

The third aspect of vocabulary knowledge in this thesis concerns vocabulary size. Schmitt (2010) believes that learners need "a lot" of words to communicate effectively in a language (p.6). However, in Schmitt (2008), it is observed that determining the necessary vocabulary size for learners will rely on the language proficiency that they wish to achieve, namely their target vocabulary. For instance, if the learners' aim is to attain native speaker level, they will need a large vocabulary size of between 16,000 and 20,000 family words (Nation & Waring, 1997; Schmitt, 2010). Nation (2013) argues that for long-term learning, second language (L2)

learners need an extensive vocabulary, but this is not necessary for a short-term goal, where learners merely need to be able to identify the words that will be useful for them to know.

To ascertain the number of vocabulary items required by learners to communicate well in their target language, Nation (2013) delineates three essential considerations when setting a vocabulary learning goal, especially on a long-term course. These key aspects consist of the total volume of vocabulary in the target language, the size of vocabulary commonly expected amongst native speakers of the target language, and the breadth of vocabulary necessary for effectively developing and using each language.

Based on earlier research, Nation and Waring (1997) estimated the vocabulary size of English native speakers as around 20,000 words. On starting school, a child's vocabulary is likely to consist of 4,000 to 5,000 words, with 1,000 word families being added to their vocabulary knowledge each year. Schmitt (2010) approximates the vocabulary size of educated English native speakers as between 16,000 and 20,000 word families, although the vocabulary size of educated native speakers in other languages, such as Arabic, is likely to be higher. Referring to Masrai and Milton's (2016) study, educated native speakers of Arabic possess knowledge of approximately 25,000 words, based on English equivalent vocabulary tests, such as Nation and Beglar's (2007) Vocabulary Size Test (VST). Masrai and Milton used the Arabic Web-based Corpus (Sharoff, 2006) as the source for developing their Arabic VST: Arabic-Lex.

By using the Wellington Corpus of Spoken English, which contains 1,000,000 running words of written British English, Nation (2006) counted a list of 6,000 to 7,000 word families required by L2 learners to be able to communicate well on an everyday basis. This corpus is obtainable from the International Computer Archive of Modern and Medieval English (Nation, 2006). However, Kuiper (2009) remarks that L2 speakers will not be capable of communicating in all genres (such as local, general and academic language) and spoken situations with native-like proficiency, "because no native speaker has native-like competence and communicative performance in all genres" (cited in Schmitt, 2010, p.7). It is important to mention that the genre used in the ELC at the College under study is general English for the Foundation Programme (FP) and academic English for the Post-Foundation programme. In the FP, which is the focus of this thesis, students are given a list of vocabularies that are

elicited from their curriculum (general English). This curriculum (see subsection 1.2.2) is provided and selected by a specialized panel in the Ministry of Manpower for all technological colleges in Oman, including this Technological College.

Furthermore, using data from the British National Corpus, Nation (2013) estimated 8,000 to 10,000 word families, including proper nouns, to reach 98% coverage of the reading of written texts, such as novels, newspapers and children's writing, with 4,000 word families enabling learners to achieve 95% coverage of written text. In contrast, for spoken English, learners need 3,000 word families to comprehend 95% of spoken text and between 6,000 and 7,000 to achieve 98% coverage. These data reflect the importance of increasing the vocabulary size for learning English language.

To contextualize the vocabulary size within the Omani context, students in Oman who come to the college, after 12 years of studying English, should be able to have at least 6,000 words, as a typical level of competency after this period of time, based on the MOE expectations (see subsection 1.3). This typical level of competency means that students can communicate well and comprehend 97% of an authentic reading text (Nation, 2013). However, Omani students at this stage only have 2,000 words according to Laufer (2000), which is much lower than native speakers at the same stage, who have an average of 18,000 words (Schmitt, 2010), and even well below the essential requirement of 10,000 words for the L2 learners to comprehend the university texts (Averianova, 2015). Therefore, the thesis aims to explore the method of vocabulary learning to increase students' vocabulary size, as the research demonstrates that the growth in the vocabulary size is crucial for language learning.

Aside from this, based on the level of vocabulary frequency, there are three frequency-based word lists, low-frequency words, mid-frequency words and high-frequency words (Nation, 2013). Nation (2013) categorises the most frequent level as 2,000 word families, whereas Schmitt (2010) goes higher to 3,000 word families for this category. Referring to Nation (2013), learners are frequently exposed to high frequency vocabulary and this is commonly repeated in text. Such vocabulary contains function words, like articles (*a*, *an*, *the*), prepositions (*in*, *on*, *for*), pronouns (*she*, *he*, *they*), and so on. In addition to function words, this type also includes content words (for example, *government*, *production*) (Nation, 2013, p.18). Thus, not much attention is required for these words, as they are easily acquired.

Meanwhile, mid-frequency words range from 6,000 to 7,000 word families (for example, *hither*, *aired*, *zoned*). Both high and mid-frequency vocabulary enables learners to communicate easily in English. Nevertheless, low-frequency words, comprising technical terminology that can be found in various fields of academic study are counted as no more than 9,000 words (for example, *bureaucracy*, *abrasion*). Words from this group are infrequently encountered in common language use and represent a very low percentage of ordinary written text (Nation, 2013).

Schmitt (2010) states that even though the vocabulary sizes cited above can weigh heavily on those working in language fields when learning and teaching vocabulary, language professionals often fail to fully recognise the real challenges faced in vocabulary learning, whereby one word family will point to a number of other words that derive or are inflected from the main root form. Therefore, both learners and teachers are left attempting to deal with 10,000 individual words. For further illustration of this point, a word family incorporates a number of individual words, such as the root, inflections and derivations; for example: *create* (root), *created*, *creating* and *creates* (inflections), and *creation* and *creative* (derivations). Nation (2006) indicates in his British National Corpus that the high frequency 1,000 word families have an average of six individual word forms per family, and an average of three words for each family word at the 9,000 word frequency level. Based on Nation's calculation, 8,000 family words are required for reading a wide range of materials, involving 34,660 individual words. Consequently, English learners must acquire an extensive vocabulary to enable them to use the English language effectively (Schmitt, 2010).

The above discussion therefore attests that vocabulary size is fundamental to learning a language and what is more, learners are required to learn a great quantity of vocabulary. Therefore, this thesis is focusing on this aspect of language learning, i.e. vocabulary acquisition, for lower level students at college. The discussion also highlights that it is important for teachers to set a long-term vocabulary learning goal, based on the learners' vocabulary size and the vocabulary frequency necessary for the targeted language proficiency level. As a result, it is appropriate here to shed some light on methods and theories of vocabulary learning, which are explored in the following sections.

2.3 Vocabulary Learning

The above section has examined vocabulary knowledge and explored the various ways in which this is measured and assessed. Beyond this, however, a closer inspection of how vocabulary is actually learned is required. For example, Hunt and Beglar (2005) consider vocabulary to be the heart of language production and comprehension. Marttinen (2008) adds that vocabulary is involved in the four language skills: listening, speaking, writing and reading. Similarly, Huyen and Nga (2003) note that vocabulary is an element connecting the four language skills: a learner cannot approach a language properly without sufficient breadth of vocabulary; it is what enables learners to communicate effectively in a foreign language. To clarify this further, a wider vocabulary helps learners to encounter fewer difficulties in their target language, resulting in better understanding and improved performance, in terms of actual use of the target language. Marttinen (2008) states that "the main object of language learning has usually been seen as achieving communicative competence" (p.14). The successful use of language to communicate can thus be achieved by enhancing methods of vocabulary learning and teaching.

In recent times, numerous theories have been introduced in vocabulary learning and teaching (Nation, 2013; Schmitt, 2010). However, the actual process of vocabulary learning is considered to be a hard task for learners (Lees, 2015) and there are consequently many different ways of approaching it. Some of the theories, relevant to this study, which underpin these approaches are discussed briefly below.

2.3.1 Theories of Vocabulary Learning

Ma and Kelly (2006) hold that vocabulary learning approaches can be divided into two main paradigms: implicit ('incidental') and explicit ('intentional') learning. Likewise, Schmitt (2005) states that there are two main theories of vocabulary learning, namely implicit ('incidental') and explicit ('deliberate'). The implicit approach is associated with meaning in an ordinary context and involves less effort, whereas the explicit approach is linked with deliberately exerted effort to connect word forms to their meanings (Ma & Kelly, 2006). The following section will explore explicit learning in detail, as this reflects the nature of the intervention used in this study.

Researchers who support explicit vocabulary teaching encourage the use of direct instructional methods, such as the use of flashcards and dictionaries (Coady, 1993; Nation, 2001). Nation and Meara (2010) add that explicit learning is effective for gaining essential vocabulary in the earliest phases of language learning and for developing language acquisition, as the learners will not have a sufficient amount of vocabulary knowledge to engage in other learning tasks (Ma & Kelly, 2006). Meanwhile, Nation (2001) advises the adoption of different approaches to teaching and learning low frequency words (the 9,000 word level), as opposed to high frequency words (the 2,000 word level). Nation considers that high-frequency words are important and should therefore be taught and learned at very early stages of language learning, via direct instruction. Therefore, the present thesis also supports an explicit vocabulary learning approach, acknowledging the importance of intentionally conveying new words to students at an early stage of their English language proficiency, given that students with a low level of English are the target sample in this study (see subsection, 3.4.2). To further explain the researcher's reasons for adopting this position, more details of explicit vocabulary learning theory are set out in the following subsection.

2.3.1.1 An Explicit Vocabulary Learning Theory

Very early on in his research, Nation (1974) outlined three main factors to be considered by teachers when teaching vocabulary: form, the meaning of the word, and the need to teach these two elements together. Also highlighted was the importance of learners having an interest in learning vocabulary, which would motivate them to exert appropriate effort in extending their vocabulary knowledge. These findings, therefore, laid the foundation for subsequent studies and their implementation in practice.

To expand on the above in further detail, Nation (1974) stated in the same study that teachers should keep their students enthusiastic about learning vocabulary through the application of various vocabulary learning techniques, involving the repetition of sequences to facilitate understanding of their meaning and at the same time, assisting students to memorise their form. It is essential for teachers to utilise methods that are suitable for their learners' level of competence and knowledge. The selected methods or tools should also be used to make the classes effective and interesting for learners (Kwiatkowska, 2007).

For instance, using an L2 vocabulary list is one of the common explicit learning strategies, but it has recently been considered unfashionable by L2 teachers (Van Benthuysen, 2003), "because teachers have not considered how it can be incorporated into larger, recursive vocabulary learning programmes" (Fernández, 2001, p.79). However, it is important to include low frequency words in lists, as Mehring (2005) remarks that "low frequency words are the more difficult aspect for teachers to teach and learners to understand… Developing word lists from the context of the lesson can reduce the work load" (p.4) for both teachers and learners, whereby learners can become exposed to new words over repeated occasions in all language skills, and therefore, grow accustomed to them. Consequently, learners also need to be provided with vocabulary lists, derived from their subject matter. Providing students with vocabulary lists or glossaries represents a strategy supported by metacognitive theory, which is explained in detail below.

A second theory relates to metacognitive strategies. According to Ellis (1995), the explicit hypothesis of vocabulary learning holds that applying several different metacognitive strategies will significantly facilitate learners' vocabulary acquisition. Metacognitive strategies include (1) noticing novel words; (2) inferring definitions through the use of dictionaries, glossaries or other assistants, and (3) promoting this new comprehension by using repetitive and associative learning strategies. Scholars have in fact found that learning vocabulary by exploiting the explicit instruction of metacognitive strategies is more useful than a teaching approach that fails to include such strategies (Rasekh & Ranjbary, 2003; Wei & Attan, 2013; Zaki & Ellis, 1999).

Wang (2000) adds to the above by stating that the use of cognitive strategies, such as note-taking, dictionaries and linguistic items, will improve learners' acquisition of new words. Therefore, links between form and meaning occur at a cognitive level of processing and through the use of metacognitive strategies (Wei & Attan, 2013). What these key theories show is that vocabulary knowledge can be developed through promoting the practice of techniques that involve noticing, sequence repetition, and associating meanings with forms. These cognitive strategies will help elicit and maintain enthusiasm amongst learners to find out the meanings of new words and recall the forms of unfamiliar words, which is also a strength of Computer-assisted Language Learning (CALL). Therefore, CALL can play a very vital function in the vocabulary learning process (Ellis, 1995; Sharifi et al., 2015).

Bearing in mind the vast number of words that are required to be learnt (see subsection 2.2.4) and the restricted time in the classes, CALL stands as an attractive tool for vocabulary learning (Miles & Kwon, 2008). Using CALL in vocabulary learning has several advantages. One of the CALL central facilities is Multimedia, in which the concepts are taught in a multi-contextual way, such as using images, text, video, interactivity and animation (Yunus et al., 2016). Secondly, a variety of techniques are used in CALL programmes to be fun- or game-based. These different techniques or CALL games, such as Spelling Games and Scrambled Words, have been designed to enhance students' enjoyment in learning through creating a challenging and entertainment environment. CALL tools allow learners to efficiently locate L1 translations and definitions of the target words. Students can also receive the immediate feedback. Computers are generally viewed as a time-saving tool. In addition, CALL tools offer students the freedom to control their learning. So, students are able to control the speed and can choose to repeat activities (Torut, 2000). According to Nation (2001), CALL can provide the basic vocabulary teaching principle, i.e. spaced repetition (see subsections 1.2.4 & 2.4.1), through offering the opportunity to optimally space practice. This has been shown to facilitate learning and retention. However, Rohrer and Pashler (2007, p.186) suggest that CALL "offers a currently unexploited opportunity to schedule study sessions in ways that optimize long-term retention". Therefore, the optimum spacing for learning L2 vocabulary is still less clear. Another advantage of CALL is to keep track of the learning (Miles & Kwon, 2008) in that teachers can track and monitor their students' progress and time spent on a certain activity or a vocabulary lesson. The track feature in CALL programmes is also important for researchers. This feature allows researchers to gather detailed data in the classroom about learning during training, along with traditional pre- and post-test data.

However, it is important to acknowledge that there are some disadvantages to using CALL. As it is a simple fact these tools are reliant on technology, students may not have the facilities at home to practise, either lacking the hardware or the access to the internet. Both teachers and learners may have negative attitudes towards using technology (Torut, 2000). So, teachers could avoid utilising technology in their classes, or students could surf something else instead of doing their work. Although there are advantages and disadvantages of CALL in theory, what is less clear is what happens in practice. Therefore, this thesis is exploring the

importance of implementing CALL in actual practice, and also investigating how CALL tools can be used to best support vocabulary learning.

2.3.2 Theories of Memory and Second Language (L2) Vocabulary Learning

The previous sections looked at the theories in learning vocabulary but this is not the whole picture. In order to enhance learners' vocabulary knowledge, the following sections will endeavour to explain how memory and the levels of processing approach can help reveal the extent to which involving CALL in a spaced repetition method will impact target vocabulary recall. Therefore, the levels of processing are a matter of importance in this study, namely understanding how it helps learners to memorise new words.

2.3.2.1 Memory and Levels of Processing

Bartlett (1932), an early pioneer of modern psychology, suggests that "memory is an effort after meaning" (p.44). To explain this more clearly, a full understanding of information will enable it to be remembered more effectively (Ellis, 2005). In fact, Bartlett's model proposes a more complex connection between sensory input and memory, as it represents the additional active process of committing something to memory. This consists of an individual's effort to memorise meaning and the form of what he or she perceives to be related to the memory of things or events experienced (Stephens, 2013).

However, this is not straightforward. Craik and Lockhart (1972) suggest further clarification of memory models, introducing the Levels of Processing Framework. In this model, information is encoded in different forms, including semantic, phonemic or visual features (i.e. verbal associations or images). Conversely, information is also processed through pattern recognition, starting with a sensory level of analysis and developing into semantic enrichment, namely forming more complex connections between conceptual representations (Ellis, 1995). Craik and Lockhart (1972, p.675) add that "memory trace persistence is a function of depth of analysis, with deeper levels of analysis associated with more elaborate, longer lasting, and stronger traces". Similarly, according to Laufer and Hulstijn (2001), the transfer of new information to the long-term memory (LTM) does not rely on the dimension of time while such information is held in the short-term memory (STM), but rather on the way it is processed, for instance shallow or deep processing. Porter (2016, p.252) also asserts that "second language vocabulary should constantly be revisited,

renewed and refreshed in order to result in longer-term learning". In this respect, deep processing is important for long-term retention, which in turn leads to detailed identification and an understanding of the memory system, as discussed below.

2.3.2.2 The Memory System

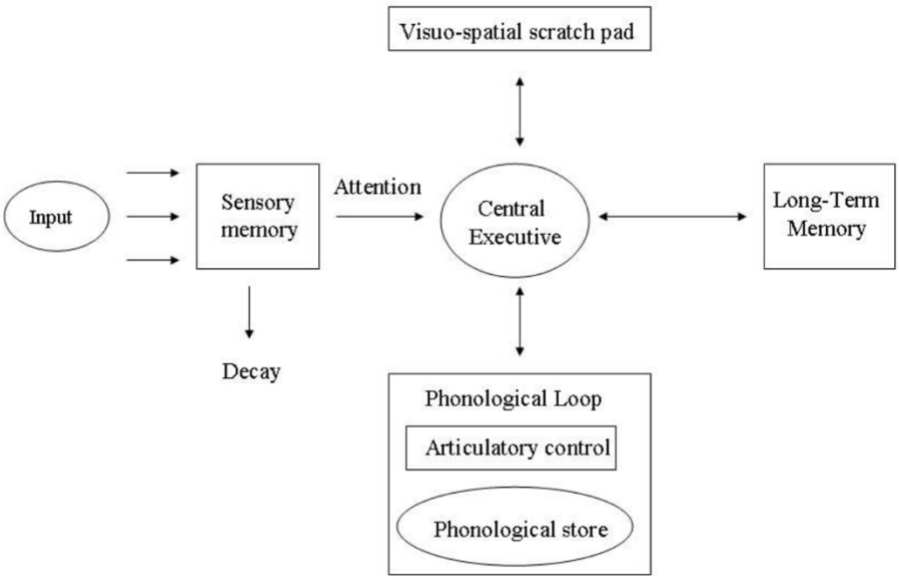
Memory is involved in almost every phase of human development throughout a person's life. It is composed of the aptitude to encode, store, manipulate and retrieve information (Ritchie, 2016). It is, therefore, helpful at this point in the research to gain an overview of existing memory models, in order to better understand how vocabulary is stored and retrieved.

Both the long-term memory (LTM) and short-term memory (STM) or working memory (WM) have been typically separated in traditional models for understanding human memory. For example, they have tended to be divided on the basis of differences in their capacity and duration (for example, Baddeley & Warrington, 1970; Craik & Lockhart, 1972; Vallar & Papagno, 1995). More specifically, the LTM is viewed as the brain's system of almost permanently storing vast amounts of knowledge. However, not all information can be directly stored in the LTM, as most will be lost from the STM. In general, however, the LTM is considered as an accessible archive of detailed information (Ritchie, 2016). In terms of a declarative, explicit form of conscious thought, the LTM is usually split into 'semantic' and 'episodic' memories. Semantic memory is a system of storing information of the meanings of words, as well as common facts about the world (Gardiner, 2001). Conversely, episodic memory is a system that enables biographical information and memorable events to be retained, but which usually declines in later years as a person ages (Craik & Rose, 2012; McLeod, 2010). In contrast, the WM is typically referred to as a system of temporary knowledge storage with limited capacity (around seven items at a time) (McLeod, 2009), and of short duration (for example, Baddeley & Hitch, 1974; Cowan, 2008; Neath & Surprenant, 2003). Although information in the WM is easily accessed and retrieved, information is also lost more easily from it as a consequence of natural deterioration (Baddeley, 2000; Burgess & Hitch, 1999; Cowan, 1999) or the role of interference (Brown & Hulme, 1995; Nairne, 1990).

It is 50 years since Atkinson and Shiffrin (1968) proposed their Multi-store Model, which separates the memory process into three stages: sensory memory (the stimulation of information), STM (the storage of information necessary for immediate processing) and LTM

(the transfer of a proportion of necessary information to permanent or near-permanent storage). Based on this model, vocabulary acquisition is linked to the growth of the LTM, while the STM is a unitary system, which is the reverse of the Working Memory Model created by Baddeley and Hitch (1974, cited in McLeod, 2008). However, Baddeley and Hitch (1974), two of the seminal researchers in this field, advocate an alternative model of STM, put forward as a Multicomponent Model of Working Memory (see Figure 2.2, below), as Atkinson and Shiffrin's (1968) Multi-store Model allows for less complexity, especially with regard to features of the STM.

Figure 2.2
Components of the Working Memory Model (Baddeley & Hitch, 1974)



In Baddeley's (1986) model, the WM includes a visuospatial sketchpad (visual information), phonological loop (verbal information), and the central executive (attentional control). In the phonological loop, information is encoded and temporarily held in the phonological store and maintained through articulatory rehearsal, i.e. repeating the information over and over (Baddeley, 2000).

In 2000, Baddeley revised his three-component model, separated into the STM and LTM systems, and proposed a fourth component, referred to as the 'episodic buffer'. The

episodic buffer works as a backup store for transferring information and communicating with different features of the memory (Baddeley, 2000; McLeod, 2008; Ritchie, 2016). However, even though the WM is not literally equivalent to the STM, it is often used interchangeably, or at least partially overlaps the STM (Cowan, 2008; Engle, 2002).

Nevertheless, despite the fact that the memory system and the terms used to explain it remain areas for disagreement among experts, the current memory frameworks generally indicate that information in the short-term storage system is lost, if there is no rehearsal, or if it is not quickly moved to the LTM (Nairne, 2002; Ritchie, 2016). This is quite an important point for language learning. Therefore, rehearsal is a key memorisation technique investigated in this thesis, because of the obvious importance in acquiring and memorising L2 vocabulary. Rehearsal or the pattern of repeating information several times helps to avoid forgetting it (Mukoyama, 2004). According to Nation (2001), it takes from 8-20 encounters with a word for it to be stored in the LTM.

However, simple rehearsal is not necessarily sufficient for ensuring that unfamiliar words are stored in the LTM in all language learners. As a result, it is also critical to consider the individual differences (Larsen-Freeman, 2014) in memory capacity. These differences among adult students are disruptive of efficiency levels in L2 vocabulary learning (Al-Hammadi, 2012). Researchers in the field of L2 acquisition have tended to give more attention to differences in STM, rather than LTM differences, in the belief that the STM represents an online capacity to process new items, such as vocabulary. Thus, the STM plays a greater role in differences in language acquisition. Based on this, an individual's greater online capacity will assist with processing new items and transferring to the off-line capacity, i.e. the LTM (Ellis, 2001).

Further to the above, according to several researchers, it is also important to establish how items are encoded at the first encounter, in order to understand how they are represented in the human mind (Greene, 1992, cited in Al-Hammadi, 2012). It is certain from the studies that the L1 has a significant impact on L2 use and learning, and this manifests in different ways (Schmitt, 2008; Swan, 1997). For instance, Nesselhauf (2003) emphasises the powerful influence of the L1 on verb-noun collocation errors, which exceeded 50% amongst advanced German learners of English in the above-mentioned study. Similarly, Hemchua and Schmitt

(2006) found that approximately 25% of lexical errors occurred due to L1 influence, in a study on the written work of EFL students in Thailand. In the last case, participants were found to engage their L1 in their L2 learning, thus explaining the excessive use of multilingual dictionaries (Schmitt, 1997, 2008). However, the best empirical studies that affirm the influential role of the L1 in L2 lexical processing and which explain the L2 acquisition process in the human mind are those conducted by psycholinguists on both L2 beginners and advanced learners (for example, Jiang, 2000, 2002; Sunderman & Kroll, 2006).

Jiang (2000, p.47) established "the first language (L1) lemma mediation stage" as the first stage of his triple-stage psycholinguistic model, based on empirical research on the role of memory in adult L2 vocabulary learning. Jiang (2002) argues that when L2 vocabulary is being processed, the L1 equivalent is activated in the learner's mind. Therefore, the use of L1 equivalent translations when learning unfamiliar L2 words is an effective means of vocabulary teaching (Ellis & Shintani, 2013). Likewise, Schmitt (2008), after reviewing several other studies, deduced that L1 translation is a reasonable way of establishing an initial connection between new word forms and their meanings.

Based on the above discussion, various means of effectively increasing vocabulary knowledge have been a subject of great interest to language learners and educational professionals alike (Tsubaki, 2007). Therefore, it was important for the intention of this research to realise the role of memory and the mechanisms for identifying L1 equivalent words in the processing of L2 vocabulary. Baddeley and Hitch (1974), who proposed the Working Memory Model, set the ground work for taking into consideration the role of short- and long-term memory in learning. Their model holds that the connection to vocabulary learning is made via rehearsal and long-term memory. The relationship between rehearsal and long-term memory supports the sequence distribution in vocabulary learning, in which this thesis seeks to find out the effect of practice distribution in the real-life setting (i.e. classroom), with low level learners in particular. However, this is not the entire story to understand the process of vocabulary learning for L2 learners. Further refinement is needed that is relevant to understanding how vocabulary is learnt efficiently. Therefore, further theories in the cognitive perspective and the levels of processing approach will be helpful to shed light on vocabulary learning, as are discussed in the following subsections.

2.3.3 The Cognitive Perspective of Vocabulary Learning

L2 vocabulary learning is considered quite a complicated process, because it not merely requires comprehension of the meanings of words, but further knowledge of how to use them productively; for example, being able to write word forms or use them in new contexts. In addition, L2 vocabulary learning involves the retention and retrieval of vocabulary (Hu & Nassaji, 2016). Here, 'retrieval' refers to the process of obtaining information (phonological form, orthographic form, meaning) about L2 words from storage in the memory (Nakata, 2016). If learners see a word form or hear a word, they thus retrieve their previous knowledge of its meaning. Therefore, Baddeley (1990) proposes that each time a word is retrieved, subsequent further retrieval is supported and facilitated (cited in Nation, 2013).

Nation and Gu (2007) state that vocabulary learning involves the remembering, processing, retention, and productive usage of L2 vocabulary. This is consistent with previous empirical studies conducted by Craik and Lockhart (1972), who formulated the concept of 'deep processing'. This proposes that the levels of processing or depth of encoding will determine the information items recalled or recognised. It was concluded from these studies that the deeper the processing, the greater the accuracy and the more likely it was for the information items to be recognised. This is quite relevant to the current study, as the levels of processing refer to vocabulary learning and the development of this notion is explained in broad terms in the following subsection.

2.3.4 Levels of Processing

The original concept of elaborative processing was proposed by Craik and Lockhart (1972) in their Levels or Depth of Processing Model. Craik and Lockhart (1972) observed that storing a new word in the LTM occurred at different degrees of encoding, i.e. the meaning of a word (a qualitative component) requires a deeper level of processing, while its form occupies a shallower level. 'Elaboration' was later added to this concept by Craik and Tulving (1975) as the key element of learning and vocabulary retention (Hu & Nassaji, 2016). However, this theory has been criticised by several scholars (Baddeley, 1978; Eysenck, 1978; Nelson, 1977), due to the ambiguity implied by measuring depth and structure in terms of levels and processing, given that there are no stable sequences of levels in shallow and deep processing.

Therefore, Laufer and Hulstijn (2001) subsequently formulated the Involvement Load Hypothesis (ILH).

According to Hulstijn and Laufer (2001), the ILH for L2 vocabulary learning is a motivational-cognitive construct of engagement based on three main components, which are used to evaluate vocabulary learning activities. These components are divided into the motivational element of involvement (need) and two cognitive elements of involvement (search and evaluation). The studies have shown that tasks with higher involvement load lead to better retention (Hulstijn & Laufer, 2001; Keating, 2008; Kim, 2008; Tang & Treffers-Daller, 2016). So, the degree of need, search and evaluation in a task will impact how well information is learnt and retained. However, the ILH does not identify which type of task (input or output) promotes more progress amongst language learners (Wong, 2017). Therefore, Nation and Webb (2011) proposed Technique Feature Analysis (TFA) to extend the three categories of the ILH, which they claim are inadequate for evaluating teaching activities.

2.3.5 Technique Feature Analysis (TFA)

TFA is a theoretical framework proposed by Nation and Webb (2011) to complement and strengthen the ILH with more criteria for the operationalisation of depth in processing. Nation and Webb (2011) proposed the TFA as a remodified initial vocabulary-learning framework, based on an earlier vocabulary learning model suggested by Nation (2001). Nation's (2001) initial framework contained three general principles of cognitive processing: noticing, retrieval and generation (Tsubaki, 2006). Noticing refers to the attention given to a new lexical set, with the aim of acquiring it, whereby learners are required to recognise words and their importance, which in turn affects their language learning process (Ellis, 1995; Nation, 2001). According to Coxhead (2008), the use of multiple vocabulary learning methods accelerates the rate at which learners 'notice' words. In this general process of vocabulary learning, the retrieval of new words occurs after noticing them. Generative use (i.e. using a target word in an original sentence or realising its other meaning or forms) then takes place after both noticing and retrieval (Tsubaki, 2006).

Nation and Webb's (2011) TFA framework introduced two further components into Nation's (2001) initial framework: 'motivation' and 'retention'. Therefore, the TFA

framework provides components that do not enrich elaboration factors, but rather suggest criteria for assessing each individual component. Thus, TFA includes a five-component framework with 18 criteria (Nation & Webb, 2011), as shown in Table 2.1.

Table 2.1
Five-component Framework of Assessment Criteria (Adopted from Nation & Webb, 2011, p.7)

Criteria	**Scores**	
Motivation		
Is there a clear vocabulary learning goal?	0	1
Does the activity motivate learning?	0	1
Do the learners select the words?	0	1
Noticing		
Does the activity focus attention on the target words?	0	1
Does the activity raise awareness of new vocabulary learning?	0	1
Does the activity involve retrieval negotiation?	0	1
Retrieval		
Does the activity involve retrieval of the word?	0	1
Is it productive retrieval?	0	1
Is it recall?	0	1
Are there multiple retrievals of each word?	0	1
Is there spacing between retrievals?	0	1
Generation		
Does the activity involve generative use?	0	1
Is it productive?	0	1
Is there marked change that involves the use of other words?	0	1
Retention		
Does the activity ensure successful linking of form and meaning?	0	1
Does the activity involve instantiation?	0	1
Does the activity involve imaging?	0	1
Does the activity avoid interference?	0	1
Maximum score		**18**

As shown in Table 2.1, the first dimension to be assessed in an activity is motivation. This identifies the clarity of the vocabulary learning goal and whether the activity will motivate learning. The ‘noticing’ criterion then highlights the attention given to new words in the activity presented, as a means of promoting learning awareness of the target word. It also involves negotiation and takes place when learners utilise a dictionary or context to get or

guess the meaning of a word, study it, or find a definition. Retrieval in this framework refers to receptive and productive retrieval, and includes recall. It evaluates whether any multiple retrievals or spacing exist between each session. Generation is then split into the receptive process (listening or reading), or productive process (using the word in a new context). Finally, retention refers to the vocabulary activity that successfully links form and meaning and ascertains whether it involves instantiation (i.e. linking a word linguistically with a meaningful example) and imaging (i.e. linking a word visually with a deliberate image), or whether it avoids interference (Farhangi et al., 2016; Nation, 2013).

As this study aims to identify effective methods of learning and retaining vocabulary via activities that link forms and meanings, it is necessary to understand how memory develops and how things are encoded in the short-term memory and then moved to long-term memory through practice, retrieving and strengthening encoding. The TFA framework can be used as a successful predictive indicator of the depth of processing facilitated by a given task (Hu & Nassaji, 2016). Hu and Nassaji (2016) advise teachers to use the TFA framework to check the features of any activities used and attempt to include more features of the TFA in these activities. For instance, having a 'generation' component is important for enhancing 'noticing' in the direction of the learner's knowledge gap (Swain, 2005). Moreover, generation provides learners with opportunities for retrieval and rehearsal, which in turn develops their vocabulary knowledge (Keating, 2008; Laufer, 2006). Nation and Webb (2011) argue that productive retrieval, illustrated in an audio-visual presentation (such as software flashcards) is strongly recommended to reinforce the mental links between form and meaning. Nakata (2008, p.5) defines rehearsal "as an activity to encode new information into our long-term memory through overt or silent articulation".

Therefore, understanding the Working Memory Model of Baddeley and Hitch and the features of the TFA framework gives a rationale to adopt the research method of explicit vocabulary learning, which facilitates the opportunity of vocabulary practice, rehearsal and retrieval to deepen the levels of processing within short- and long-term memory.

2.4 Vocabulary Teaching and Practice

The previous sections in this literature review have looked at learning theory, in order to elicit and develop a better understanding of how vocabulary is learnt. The following subsections focus on the application of these theories in practice, which is the main focus of this thesis.

2.4.1 Distribution of Practice

Through extensive empirical studies, spaced learning has been favoured as an effective means of enhancing long-term memory, where information is learned over sequenced time periods, rather than a high volume of information being learned in a single session (see, for example, Cepeda et al., 2006; Lotfolahi & Salehi, 2017). Lotfolahi and Salehi (2016) found that students who spent five minutes daily learning a vocabulary list over a period of three days, performed more successfully in a retention test than another group of students, who spent 15 minutes studying the same vocabulary in a single session. This phenomenon is referred to as the 'spaced effect' and it has been a dynamic topic in the psychological sciences. In recent years, the spaced effect has been extended more widely to evaluate and test concepts and theories in areas of L2 learning (Lotfolahi & Salehi, 2017).

Schmitt (2008) recommends spaced repetition as a vocabulary-learning method for the arrangement of vocabulary learning. To date, a vast body of research on memory has revealed a preference for the positive effects of spacing in vocabulary learning, which outweigh the effects of massed repetition of words (see Kornell, 2009; Nakata & Webb, 2016; Swehla et al., 2016). Nevertheless, although spaced practice has been widely demonstrated in educational settings, it is still uncertain how it can yield benefits as a standard pedagogical methodology for vocabulary learning. It therefore constitutes one of the original contributions to knowledge made by this current study.

Several empirical studies on vocabulary learning have been carried out in educational contexts to reveal the advantages of spaced over massed practice. For instance, Kornell (2009) investigated the effect of spaced learning in three experiments, conducted using a sample of undergraduates in a realistic context, with flashcards implemented as a tool for learning 40 words. The results showed that the spaced studying of a large stack of flashcards was more effective than the massed studying of four smaller stacks of flashcards,

implemented separately. Spaced studying has also proved more effective than 'cramming' as a learning approach, i.e. studying intensively just one day before a test (Kornell, 2009). Therefore, the preferred performance in the final test was oriented towards the participants studying under spaced conditions.

In another, more recent study, Lotfolahi and Salehi (2017) implemented a novel method, using two different spacing schedules (massed versus spaced) in an English as a Foreign Language (EFL) classroom over a six-week period. Their aim was to discover the effect of spacing techniques on EFL vocabulary learning in a real-world classroom. The participants were young EFL learners, studying in an English language institute in Iran. The sample consisted of 28 students aged 7-11 years. In the first week, the participants studying under massed conditions were given five English-Farsi word pairs to learn. Further five-word pairs were presented in another session the following week. However, under spaced conditions, the participants studied all 10 word pairs in the first session, and then went over them again in a second session one week later. All the participants subsequently took two tests: one in the first week and the other in the sixth week. The results indicated that spaced practice promoted better long-term retention than massed practice. The above researchers therefore suggest spacing and retrieval practices as pedagogically powerful ways of enhancing retention and vocabulary learning, although further studies are necessary to confirm this finding. Extending this agenda, it is worth comparing the effects of different spacing intervals between practice sessions (i.e. the lag effect), which is highly relevant to what is discussed in the second language acquisition (SLA) literature, and to what occurs in typical learning settings (Rogers, 2017). Toppino and Gerbier (2014) point out the importance of longer spacing between sessions, due to its positive effect on long-term learning. However, the focus of previous SLA exploratory studies on spacing practice has been on short-term learning (Serrano & Huang, 2018), which is one of the knowledge gaps that needs to be filled in this current research.

Pavlik and Anderson (2005) propose the activation-based memory model to explain the absence of spacing effects at shorter retention intervals. The above authors state that "each time an item is practised it receives an increment of strength but that these increments decay as a power function of time" (Pavlik & Anderson, 2005, p.559). In each exposure to an item, the rate of decay is defined according to the degree of activation at the time of encountering

it. To explain this further, the activation of an item at the second encounter is high in the case of massed practice. This higher degree of activation prevents the item from being encoded in the long-term memory. Therefore, a lack of spacing between sessions leads to shorter retention time (Serrano & Huang, 2018).

Rohrer and Pashler (2007) examined the ways in which retention and learning schedules interact, finding that the length of the intersession interval (ISI) (i.e. the period between study sessions) and the retrieval interval (RI) (i.e. the period between the last treatment condition and the time of the test) had an effect on the duration of the knowledge retention. They suggest that the optimal ISI should be 10%-30% of the RI. To clarify this, Serrano and Huang (2018) present an example of a lesson on the present perfect, taught to two groups in three sessions, but presented in different ISIs. For instance, if the lesson is taught on three consecutive days (1-day ISI) and the test is conducted 10 days later, following the last training session (10-day RI), the 1-day ISI will be 10% of the RI. This is considered by Rohrer and Pashler (2007) to be optimal spacing for measuring learning. However, if the same lesson is taught once a week for three weeks, the ISI will consist of seven days between sessions, representing 70% of the RI. However, this ratio is not optimal, according to Rohrer and Pashler (see Figure 2.3). For the 7-day ISI, in this case, the optimal time at which to administer the test would be within the 30-day RI, as the 7-day ISI will be 23% of the RI, thereby keeping it within the optimal range identified by Rohrer and Pashler.

Figure 2.3
Example of Different Intersession Intervals (ISI) and Retention Intervals (RI) (Serrano & Huang, 2018, p.5)

	Session 1		Session 2		Session 3		Test
1-day ISI	Monday	1 day	Tuesday	1 day	Wednesday	10 days	ISI=10% RI
7-day ISI	Monday	7 days	Monday	7 days	Monday	10 days	ISI = 70% RI

A number of studies (for example, Bird, 2010; Rogers, 2015; Suzuki, 2017) have been undertaken to further investigate the lag effect (longer versus shorter spacing between practice sessions) in L2 learning, taking into account the interaction between the ISI and the RI proposed by Rohrer and Pashler (2007). Bird (2010) explored the interaction between two different lag effects (14-day ISI versus 3-day ISI) and two intervals (7-day RI versus 60-day

RI) on the retention of English syntax by university students in a classroom setting. The results revealed no significant differences, as the 7-day RI was not optimal for either group: 200% of RI for the distributed group (14-day ISI) and 43% of RI for the concentrated group (3-day ISI). However, the results revealed that the distributed group outperformed the concentrated group (with a large estimated effect size) after a 60-day test interval. This result was also consistent with claims made by Rohrer and Pashler, as the 14-day ISI of the distributed group was found to be optimal, at 23% of RI, while the 3-day ISI of the concentrated group was 5% of RI, which was not an optimal ratio between 10% and 30%. Similarly, the findings from Rogers' (2015) study are in line with those of Bird (2010). Both studies support that longer spacing between sessions leads to better learning and knowledge retention than shorter spacing between sessions. However, the RI spacing declared to be optimal for the delayed post-tests in both studies was only optimal for the spaced group.

Conversely, Suzuki's (2017) study, replicated from Suzuki and DeKeyser (2017), looked at shorter (3.3-day ISI) and longer (7-day ISI) spacing for the L2 morphology learning of a novel miniature language. The results supported those found by Suzuki and DeKeyser. Both the above-mentioned studies found shorter spacing to have a higher learning advantage, although there was no statistically significant difference identified between the groups. Correspondingly, Rogers and Cheung (2018) conducted a study in an authentic classroom setting to assess the impact of shorter (1-day ISI) and longer (8-day ISI) spacing between two training sessions, with regard to learning English adjectives (describing people). The sample consisted of young learners (aged 8-9 years), who were assigned two multiple-choice tests (a pre-test and a delayed post-test at 28 days). The results showed that shorter intervals between practice sessions led to better retention. However, Kasprowicz et al. (2019) found no statistically significant differences between short (3.5-day) and long (7-day) intervals. Both spacing intervals were revealed to be beneficial for young learners seeking to acquire French morphosyntactic structures. In general, mixed results have been generated by these previous studies. Further, they were all focused on grammar learning, rather than vocabulary learning. Therefore, more studies are needed on this aspect of L2 learning.

In a recent study focusing on vocabulary retention, Serrano and Huang (2018) investigated the effects of time distribution on five repetitions of a short passage, with regard to L2 fluency and vocabulary acquisition. The above study focused on the optimal lag effect,

i.e. the ideal space between reading sessions of the same text, especially to promote the maximum impact of L2 learning and vocabulary acquisition. It included 71 EFL Taiwanese students (aged 16 years), grouped into two different time distributions: intensive (1-day ISI) and spaced (7-day ISI). Two delayed post-tests were administered, as the study endeavoured to ascertain the optimal spacing range for both treatment conditions, which was an issue identified in studies by Bird (2010) and Rogers (2015). The first delayed post-test took place four days after the last session, representing 25% of RI for the intensive group. Meanwhile, in the 28-day delayed post-test, the second test was 25% of the RI for the spaced group. Therefore, both groups had the same optimal spacing. The results showed that shorter spacing led to immediate improvement in vocabulary learning. However, the longer spacing supported greater long-term retention. Therefore, this thesis seeks to mirror the design of Serrano and Huang's study to compare longer and shorter spacing distributions, but with different learners and in another context (see section 3.4).

Aside from the above, Lotfolahi and Salehi (2016) make the point that the repetition of tasks can be tedious, especially with intensive rather than spaced presentation. The latter is highly preferable, as the tasks are distributed across longer-spaced sessions. Therefore, the current study harnesses technology for various exercises in both treatment conditions (intensive and spaced) to avoid the boredom factor, with four types of receptive and productive knowledge being evaluated (Laufer & Goldstein, 2004) (see subsection 2.2.2).

The use of technology as a pedagogical method of vocabulary learning has led a significant number of researchers to integrate vocabulary software programmes for teaching and learning new words (Lin, 2010). Around two decades ago, large numbers of teachers incorporated computer programmes and the internet into their lessons, since CALL appealed to them and to educational researchers as a new outlook for the teaching and learning of vocabulary (Cellat, 2008). This phenomenon is described in more detail in the following subsection.

2.4.2 Computer-assisted Language Learning (CALL)

CALL relates to computer use for the purpose of learning or teaching a language (Sun, 2003). Similarly, Egbert (2005, p.3) simply defines CALL as "using computers to support language teaching and learning in some way". Earlier, in the 1990s, Levy defined it briefly and much

more broadly as “the search for and study of applications of the computer in language teaching and learning” (1997, p.1). This last definition is consistent with the approach adopted by most modern CALL practitioners (Davies, 2000) and the one applied in this present study.

2.4.2.1 Overview of Computer-assisted Language (CALL) Instruction and Vocabulary Learning

In the early 1980s, once CALL had been initiated and vocabulary learning had become the focus of attention amongst scholars, attempts were made to integrate vocabulary learning into CALL programmes. However, these programmes were limited in terms of the language learning activities that they offered; usually there was just one type, such as a gap-filling exercise, vocabulary game, etc. (Beatty, 2013). Ma and Kelly (2006) explained that these activities were limited in their variety, not only because of the limitations of computers at the time and the lack of expert teachers in the field of programming (Levy, 1997), but also due to the dearth of vocabulary learning theories. All these factors hindered the development of effective vocabulary learning programmes (Ma & Kelly, 2006).

However, computer technology has moved on since Ma and Kelly’s (2006) study and even prior to it, Moras (2001) emphasised that computers were already in widespread use. Thus, anyone who did not use a computer was likely to feel outdated. In this era of ubiquitous technology use, a range of computer programmes has been introduced into the teaching and learning of L2 vocabulary, providing learners with new and enjoyable ways of applying their skills through audio-visual resources (for example, Anki, Word Engine and Quizlet) (Dizon, 2016). The current enthusiasm for much greater pedagogical use of CALL has emerged as a consequence of the rapid increase and modern adaptation of language software (Bagheri et al., 2012).

Nowadays, CALL goes far beyond traditional paper and pencil methods, offering various types of activity for practising and learning a language (Özer & Koçoğlu, 2017). For example, Kilickaya and Krajk (2010) investigated Turkish undergraduate students, using an online vocabulary programme entitled WordChamp. The results showed that online vocabulary learners performed better than learners who studied vocabulary using traditional teaching and learning methods. A delayed post-test was administered two months after the post-test, as a follow-up investigation of the learners’ long-term vocabulary retention. The

results demonstrated that the experimental group had retained more vocabulary knowledge than the control group, with 4% and 6%, respectively, of the target words being lost.

Likewise, Nakata (2011) asserts the advantages of using the iKnow CALL programme, which offers opportunities for practising different types of activity, such as multiple-choice questions and other tasks to promote both receptive and productive vocabulary knowledge. In addition, Nakata's (2011) study investigated the short- and long-term effectiveness of two vocabulary learning conditions (CALL with spaced repetition versus vocabulary notebooks). The results revealed that both conditions produced equally satisfactory results, with no differences in the short term. Conversely, the longer-term results favoured the CALL group, with 5% loss, while the vocabulary notebook group were found to have lost 14% of the target words (cited in Özer & Koçoğlu, 2017).

The above results correspond to Al-Khatib's (2011) study, which demonstrated the positive influence of this technology on support for language learning. Al-Khatib conducted a case study on 43 English language and literature students in the Arab Open University in the Lebanon, using technology to facilitate the learning of eight different themes related to English language learning. The participants were divided into eight groups according to the theme that they were working on collectively. The findings revealed that technology use led to higher performance in the final test. Therefore, Al-Khatib confirmed the positive effect of technology use on the learning process and supported the learner-centred experience.

2.4.2.2 Overview of the Literature on Flashcard Software Programmes

The current study aims to explore whether or not using a flashcard software programme is effective for promoting vocabulary learning. First, flashcard programmes or computer-based flashcards are referred to as software that supports L2 learners in learning target words, presented as a paired-associate task. To clarify this further, such programmes introduce the target words far more comprehensively than meaning-focused activities, asking students to link word forms with their meanings (i.e. L1 equivalent translation and L2 synonyms or definitions) (Nakata, 2011). Several types of software flashcard have been used extensively. For example, the programme vTrain (http://www.vtrain. net/) has been implemented by several hundred schools and 50 universities worldwide (Rädle, 2009), whereas the number of active users of the Quizlet flashcard programme (http://quizlet.com/) exceeds 30 million each

month (Quizlet, 2018).

2.4.2.3 Overview of the Literature on Quizlet

The main tool forming the focus of this study is Quizlet. Therefore, it is important to spend some time reviewing the studies that have looked at Quizlet (for more explanation of the programme's functioning and the rationale for choosing this software, see subsection 3.4.4.8). Recent studies on support for vocabulary learning have used Quizlet to stimulate vocabulary acquisition (Anjaniputra & Salsabila, 2018; Chien, 2015; Jackson III, 2015). However, these studies have mainly focused on students' perceptions of its use. They are therefore limited and fail to take several factors into consideration. For example, they have not applied any vocabulary theories on the use of software, and there are no achievement test reports (i.e. on pre-, post-, or delayed post-tests). Moreover, they have not taken into account the factor of word length or parts of speech. In addition, in the current generation of flashcard software programmes, such as Anki, Word Engine and VocabTutor, which have generally been considered as practical tools for expanding learners' ability to acquire new words, Quizlet has stood out as an interactive device that empirically facilitates vocabulary learning (Dizon, 2016). Nevertheless, although it is widely used for this purpose, few research studies have been conducted to investigate the influence of Quizlet's implementation on learning vocabulary items (Crandell, 2017).

Meanwhile, Chien (2015) affirms the significant impact of flashcard use to present and practise new words in the classroom, utilising Nakata's (2011) 'Criteria for the Evaluation of Flashcard Software' and Nation's (1994) 'Activities for Vocabulary' to evaluate three online flashcard websites: Flashcard Exchange (Cram.com), Study Stack, and Quizlet. The findings from this comparison, using a checklist, indicate that all these programmes enabled the students to acquire both productive and receptive vocabulary knowledge, and offered presentation and retrieval modes. Chien subsequently found that the focus of the three websites was essentially on form (spoken and written) and word meaning (connecting form to meaning), based on Nation's (1994) learning model for vocabulary activities, as illustrated in Table 2.2. Moreover, the exercises on these websites matched eight of the criteria presented by Nakata (2011) (see Table 2.3). However, Chien did not measure students' vocabulary knowledge, or explore the impact of the three online flashcard websites on the learners' vocabulary knowledge using achievement tests, such as pre-, post-, and delayed post-tests.

Table 2.2

Comparison and Contrast between Online Vocabulary Flashcard Websites Based on Nation's (1994) Vocabulary Learning Activities

Goal	Elements	Quizlet	Study Stack	Flashcard Exchange
Form	Spoken form	√		
	Written form	√	√	√
	Word parts			
	Form-meaning	√	√	√
	Connection			
	Concept and references			
Meaning	Association			
	Grammar			
Use	Collocations			
	Constraints on use			

Table 2.3

Comparisons between Three Online Vocabulary Flashcard Websites Based on Nakata's (2011) Criteria

Criteria	Quizlet	Study Stack	Flashcard Exchange
I. Flashcard creating and editing			
1. Flashcard creation	√	√	√
2. Multilingual support	√	√	√
3. Multiple words			
4. Types of information			
5. Support for data entry	√	√	√
6. Flashcard set	√	√	√
II. Learning			
1. Presentation mode	√	√	√
2. Retrieval mode	√	√	√
3. Receptive recall			
4. Receptive recognition			
5. Productive recall	√	√	√
6. Productive recognition	√	√	√
7. Increasing retrieval effort			
8. Generative use			
9. Block size			
10. Adaptive sequencing			
11. Expanded rehearsal			

Chien (2015) also recorded observations and conducted group interviews among 20 randomly selected participants in their freshman year at a Taiwanese university, in order to find out their opinions of the three above-mentioned online programmes. In her qualitative findings, the participants showed a strong preference for Quizlet, especially the 'Speller' activity (see Figure 2.4), followed by the 'Test and Space Race' (Gravity) activities. The above researcher subsequently recommended further interventions to compare flashcard websites and include learners with diverse levels of English language proficiency.

Figure 2.4
Speller Activity on the Quizlet Website

Similarly, in a recent study, Anjaniputra and Salsabila (2018) illustrated the perceptions of 30 tertiary level students on the use of Quizlet to learn complex vocabulary. The above authors used two qualitative methods (observation and interview) to collect data from the participants. The findings showed that Quizlet enhanced the learners' engagement and tenacity to learn vocabulary, and the learners found Quizlet enjoyable and useful for their vocabulary learning. These results support the findings from Lander's (2016) study, conducted on 830 low-intermediate level students. Lander investigated these students' views after nine months of using Quizlet, both within and outside the classroom. The results showed that the students valued Quizlet and considered it to be a new and enjoyable means of learning vocabulary.

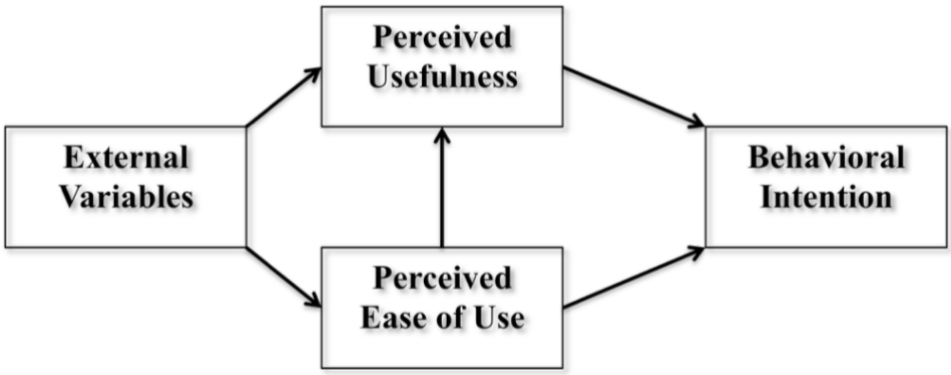
Further to the above, Jackson III (2015) found that Quizlet elicited more favourable views than Educreations, in a study where instructional videos were created by the teacher and shared with students through an application on the students' mobile phones. The study involved 54 female university students on a Foundation Programme in Al Ain, who took part in a study to investigate the motivating role of L1 in L2 vocabulary acquisition. For this purpose, two mobile learning applications were used: Quizlet (in the home for 10-15 minutes per day, practised through vocabulary exercises, such as L1 translation) and Educreations (in class, watching 5-7 minutes of instructional videos, followed by activities that included L1 translation). Jackson III (2015) presented three main reasons for favouring Quizlet, namely the fact that a grade was received directly after completing each session, the matching of L2 vocabulary with its L1 equivalent, and the inclusion of games. Most of the participants were found to be enthusiastic about using Quizlet for the allocated time (10-15 minutes), based on its benefits in the form of extra exercises outside the classroom. They also agreed on the importance of Arabic translation to accelerate the rate at which they assimilated new vocabulary. However, as the researcher deliberately selected exceptional female students, who had previously attained high vocabulary scores in a computer-generated exam (CGE), the results cannot be generalised to male learners, or to learners of either gender with lower proficiency.

Correspondingly, Dizon (2016) conducted a study to explore the efficacy of using Quizlet flashcard software to develop L2 English vocabulary. His sample consisted of nine Japanese university students in their second year at the Faculty of Foreign Studies in Japan. These participants were selected based on their TOEIC exam scores from the previous academic year, thus ensuring that they were among the top English L2 students. The above-mentioned vocabulary course lasted for 10 weeks at a frequency of three sessions per week, covering 570 terms from Coxhead's (2000) Academic Word List (AWL). Students were given a pre-test using Version One of the 30-item vocabulary level test (VLT) and Version Two of the VLT as a post-test.

Dizon developed a questionnaire based on the influential Technology Acceptance Model (TAM) presented in Davis's (1989) research framework (see below), in order to measure the learners' behavioural intention (BI) to use technology (Quizlet). The questionnaire examined two main technology factors, namely perceived usefulness (PU) and

perceived ease of use (PEOU). Both PU and PEOU play a role in defining users' BI, as well as considering the involvement of certain other external variables. Accordingly, the TAM framework contains three main components, as shown in Figure 2.5

Figure 2.5
The Technology Acceptance Model (TAM) (Davis, 1989)



Davis (1989) defined PU as "the degree to which a person believes that using a particular system would enhance his or her job performance" (p.320), whilst PEOU refers to "the degree to which a person believes that using a particular system would be free of effort" (p.320).

The results of the above study indicate Quizlet as a useful online tool for significantly improving participants' VLT scores, and that the participants had positive perceptions of using it, with mean scores that were higher than '4' on a 5-point Likert scale for all three components of TAM (PU, PEOU and BI). The participants also appeared to favour using their mobile phones, thus demonstrating a preference for mobile learning technology. Consequently, Dizon strongly advocated the use of Quizlet in the classroom. However, the sample size of just nine participants was very small. In addition, no delayed post-test was implemented as a measure to check the learners' long-term memory and no control group was included in the study.

Along similar lines, Özer and Koçoğlu (2017) assessed the effectiveness of two methods of vocabulary learning: Quizlet and a vocabulary notebook, using a larger sample in a Turkish vocational High School. The sample consisted of 89 participants, drawn from four classes: two classes serving as the experimental groups (Quizlet and Notebook), and the other

two operating as control groups. The findings of Özer and Koçoğlu's study showed that both experimental groups significantly outperformed the control groups in the post- and delayed post-test. They mentioned that the Quizlet group had made slightly more improvement than the Notebook group. It is worth mentioning here that to the best of the present researcher's knowledge, Özer and Koçoğlu's study is the only quasi-experimental study to have attempted to measure Quizlet's effectiveness for long-term retention via a delayed post-test. However, the delayed post-test took place just two weeks later, which cannot be considered sufficient for assessing the learners' long-term vocabulary retention.

Other recent research (Korlu & Mede, 2018; Zambrano Acosta, 2018) has explored the impact of using Quizlet by collecting both qualitative (participants' perceptions) and quantitative data (from performance tests, i.e. pre- and post-tests). However, these studies have also omitted to implement a delayed post-test. In this regard, Zambrano Acosta (2018) implemented a quasi-experimental design to investigate Quizlet's effectiveness for vocabulary learning amongst teenagers with a low level of English language ability. The participants comprised 50 students, divided into an experimental group (N=23) and a control group (N=27). The experimental group learned the vocabulary via the Quizlet App over a period of four weeks, whereas the control group were assigned to learn vocabulary via traditional means, such as using textbook activities and keeping a vocabulary notebook. Zambrano Acosta employed a pre-test before the intervention, followed by a survey, with a post-test administered one week later. The pre- and post-tests were focused on the listening and reading activities, with results that showed a noticeable improvement in the experimental group's vocabulary knowledge and English language performance. Conversely, the control group showed no change or improvement. The survey findings revealed that the students had positive views of using technology in general and Quizlet in particular to learn vocabulary.

Similarly, Korlu and Mede (2018) adopted a quasi-experimental research design to evaluate 40 Turkish pre-intermediate EFL university students, with regard to their performance and autonomy in vocabulary learning. These participants were drawn from two intact classes: one class was recruited as an experimental group who would use Quizlet, and the other was assigned as a control group, who would use a vocabulary notebook. Korlu and Mede also investigated the participants' views of incorporating Quizlet as a mobile learning application into English language preparatory classes. To collect these qualitative data, Korlu

and Mede distributed an online survey to the students, conducted semi-structured interviews with 15 of the students, and assigned a reflective journal to the teacher. The findings from a Friedman's ANOVA test showed that both groups had significantly improved by the post-test, following eight weeks of treatment. However, the descriptive statistics revealed that the difference between the average scores in the experimental group's pre- and post-tests was far higher (53.69 points) than that of the control group's average grades in these tests (32.11 points). Therefore, Korlu and Mede concluded that Quizlet was more effective for learning and teaching vocabulary. In addition, both the students and their teacher expressed positive views of using Quizlet. Thus, Korlu and Mede suggest integrating Quizlet as a mobile learning application into English preparation classes.

Furthermore, Tam et al. (2010) encouraged the use of CALL programmes amongst learners with low levels of English language proficiency. They state that CALL programmes help create an ideal environment for such learners, because they increase their confidence and motivation (Krish et al., 2011). However, Barr (2016) and Sanosi (2018) are amongst the very few researchers to investigate low-proficiency English learners at university level, with a view to enhancing their vocabulary knowledge through Quizlet. Barr (2016) used a sample of 32 first-year Japanese students, who accessed Quizlet to prepare for vocabulary tests, re-visiting gap-filling content from the Quizlet flashcard sets that were provided in class. The students were given instructions on how to use Quizlet and were encouraged to do so. They were then classified according to their individual study habits. The resulting scores showed that the Quizlet users outperformed the non-users in the same class. With new content, the Quizlet users also recorded moderately higher scores. Barr recommends further research on a more homogeneous sample, assigning activities that involve spaced repetition.

In a similar vein, Sanosi (2018) adopted an experimental research design to investigate the effect of Quizlet on vocabulary acquisition. The participants comprised 42 EFL beginners in their first year at a Saudi university. This sample was randomly allocated to two groups: an experimental group (taking regular classes and using Quizlet within and outside the university) and a control group (who only took regular classes). The study instruments consisted of pre- and post-tests. The results of the t-test showed no significant differences between the groups in the pre-test. However, the experimental group appeared to have significantly progressed and outperformed the control group by the post-test. In contrast, the

control group showed no significant change in either the pre- or post-test. In the above study, Sanosi considered Quizlet to be an ideal tool for learning vocabulary, whether within or outside the classroom. However, his study is yet another piece of research that fails to include a delayed post-test to assess the learners' long-term recall. Moreover, there was no control over the amount of exposure to the target vocabulary or types of activity allocated to the students during the intervention, because they were also allowed to access Quizlet outside the university.

A wide range of recent studies have emphasised the advantages of using software to help students learn vocabulary. These studies have generally favoured Quizlet, based on students' results and perceptions. This preference has also resulted from the application of specific criteria, following comparisons between Quizlet and other software, and compliance with standards such as those outlined by Nation (1994) and Nakata (2011) (see Chien, 2015). Previous studies on Quizlet have only looked at learners' perceptions of these tools, without investigating how Quizlet improves productive and receptive knowledge by conducting experimental studies to establish what has been learnt. Moreover, in addition to learners' perceptions and technology use in actual practice, it was considered essential in this current study to benefit from teachers' perceptions and experiential knowledge, as teachers "are the best people to design the courses they teach" (Graves, 2000, p.5).

2.4.2.4 A Brief Overview of Teachers' Perceptions of the Use of Computer-assisted Language Learning (CALL)

A vast body of literature supports the positive impact of technology use on language learning (for example, Afrin 2014; Al-Khatib, 2011; Hani, 2014). Al-Khatib (2011) points to the positive influence of technology use on the significant enhancement of learners' achievements in a final exam and an increase in their ability to comprehend, analyse and engage in dialogue, as well as gaining social-cultural and academic benefits. In addition, Afrin (2014) states that teachers generally have positive attitudes to the integration of CALL into their classes. Afrin suggests that technology helps learners take responsibility for their own learning, enabling them to be independent and build confidence in their learning, so that they interact in a freer atmosphere, where meaningful and authentic learning can take place. Similarly, Hani (2014) specifies several benefits of CALL for EFL learners, indicating that technology helps to integrate language skills (writing, reading, listening and speaking), both

in and outside the classroom. As a result, immediate feedback can be provided, individual work promoted, motivation and fun increased, and time and effort saved.

Surrounding the topic of vocabulary learning, which is what this thesis mainly addresses, Lee (2005) revealed the effects of CALL on improving vocabulary by pointing learners towards the use of an online dictionary. Lee subsequently found that learners also improved their grammar, held positive perceptions of using CALL, and expressed an intention to use it in future. This was likewise found in Zapata and Sagarra's (2007) study, in which a large number of learners held positive views of CALL's impact on their pronunciation and vocabulary. Correspondingly, Lin et al. (2008) demonstrated that CALL helped enhance EFL learners' achievement in pronunciation and vocabulary, as well as their speaking and reading skills. Therefore, the results of previous studies emphasise the important role of CALL in improving vocabulary knowledge.

In the context of Oman, there is a significant push towards integrating technology into higher education institutes (HEIs) (see subsection 1.2.2). However, relatively few studies have investigated teachers' perceptions of embracing technology in language learning at Omani HEIs (Al-Musawi, 2007; Al-Musawi & Abdalraheem, 2004; Al-Senaidi et al., 2009; Ambu-Saidi, 2010). These studies have primarily reported perceptions of technology use in language learning. For example, Al-Musawi and Abdalraheem (2004) conducted their study in Sultan Qaboos University (SQU), the first government university in Oman, finding that both teachers and students had positive perceptions of using technology on courses, as it increased their self-confidence in solving problems. However, the teachers and students cited a technical problem, relating to internet speed. Moreover, the teachers considered that it was time-consuming to use technology for learning, because of the amount of time required to design online learning materials, download assignments, express ideas, and respond to emails. Irrespective of all these human issues and technical obstacles, Musawi and Abdalraheem claimed that the implementation of technology in HEIs is promising, indicating a rapid increase, with a strong focus on implementing technology and digital literacy on the part of the Omani government. Ambu-Saidi (2010) also noted a preference amongst teachers and students from an Omani College of Applied Sciences to use technology, as it allowed them to access a variety of resources. It also motivated the users and offered them some flexibility with regard to time and place.

In contrast, one year before Ambu-Saidi's study, Al-Senaidi et al. (2009) conducted research in another College of Applied Sciences, revealing negative perceptions of technology use in teaching. Their study involved 100 teachers from four different departments at the College. The teachers listed several barriers to the adoption of technology, including "lack of equipment, lack of institutional support, disbelief of ICT benefits, lack of confidence, and lack of time" (Al-Senaidi et al., 2009, p.575). Likewise, in a previous study, Al-Musawi (2007) referred to negative views of technology integration into the learning process, attributing these to low awareness of the positive impact of technology use on language learning. Therefore, Al-Musawi emphasised the importance of ongoing assistance to support the integration and use of technology by administrative and academic staff, and the need to raise their awareness of technology's importance for promoting language learning. This is supported by the TPACK model, which emphasises the importance of combining technological, pedagogical and content knowledge for the effective integration of technology into the classroom (Koehler & Mishra, 2009). This model, developed by Mishra and Koehler (2006), considers technology awareness and knowledge to be a fundamental domain in the Technology, Pedagogy, and Content Knowledge framework (TPACK). Furthermore, Kim (2002) asserts that teachers' positive perceptions of technology also reflect positively on learners' knowledge and implementation of technology. This underlines the importance of seeking teachers' opinions in this thesis, as they are in charge of the classroom.

Therefore, the literature reviewed in this chapter helped the researcher to identify a gap in the research and formulate research questions for this study (see section 3.2). The following section will now summarise the points that emerged in the course of the review and conclude this chapter.

2.5 Summary

This chapter has addressed vocabulary knowledge, vocabulary learning and vocabulary learning theories, with attention largely being given to some of the issues surrounding vocabulary development in terms of word knowledge, word counts and vocabulary size. In addition, the notions of productive and receptive knowledge were introduced. Finally, the effectiveness of spaced repetition for vocabulary retention was illustrated, this being the method applied in the present study.

Based on a review of the current literature, linked with theories of memory and vocabulary learning, several gaps were revealed in this chapter. As a result, two main areas of investigation were highlighted: first, effective ways of acquiring vocabulary for the long term, given that the number of studies dedicated to long-term memory in second language acquisition (SLA) is very limited. This area was selected, due to the fact that word memory is thought to be the most likely determinant of language development (Ellis, 2001). In addition, there is a lack of experimental studies on vocabulary acquisition that use Quizlet in real-life settings. Therefore, the present thesis sets out to examine the outcomes of learners with low levels of English proficiency by examining performance in achievement tests (pre-, immediate post-, and delayed post-tests). Thus, the results of the intervention could explain and fill a knowledge gap in vocabulary learning by measuring the learners' short- and long-term memory. Secondly, the effectiveness of intensive versus spaced distribution practice in vocabulary acquisition will be explored by assessing learners' vocabulary knowledge, following both short and long instructional periods (i.e. longer versus shorter spacing between sessions). Therefore, this thesis seeks to fill the knowledge gap, concerning the impact of the lag effect, by examining the performance of low-proficiency learners in a real-life setting (i.e. the classroom), rather than a clinical study setting (i.e. a laboratory).

This chapter has also reviewed vocabulary learning and CALL, giving an overview of the literature on flashcard software programmes and teachers' perceptions of the use of CALL. In reference to Jiang's (2000) model, learners will consequently be provided with L1 equivalent translations to facilitate L2 vocabulary acquisition. Therefore, the present study integrates a flashcard software programme, enabling learners to obtain the L1 equivalent of a word by clicking on the L2 word or checking the vocabulary list provided by the programme. However, Ma (2013) emphasises the importance of laying down a foundation of pedagogical theory when designing Computer-assisted Vocabulary Learning (CAVL) programmes. Therefore, the present study adopts the TFA and TAM frameworks as its theoretical bases for the selection and evaluation of the software programme deployed, and to ascertain the effect of the latter on vocabulary learning.

CHAPTER THREE: METHODOLOGY

3.1 Introduction

This Methodology chapter addresses the research questions and hypotheses, followed by a discussion of the philosophical and methodological assumptions that are adopted in the present study. It also justifies the research design and underpinning theory, describing the research context, participants, data collection, and data analysis techniques applied. In addition, it outlines the reliability, validity and ethical considerations of this research, concluding with a summary.

3.2 Research Questions and Hypotheses

This study addresses three main questions:

1. Is Quizlet (a Computer-assisted Language Learning [CALL] tool) an effective programme for promoting vocabulary learning amongst low ability learners in the classroom?

2. Does the time distribution of the practice sessions (intensive versus spaced) moderate the benefits of using Quizlet to promote vocabulary learning and retention amongst low ability learners?

3. What are the perceptions of teachers and students regarding the use and implementation of Quizlet?

Three hypotheses were also formulated for this research, based on a review of the literature:

Hypothesis 1.1: There will be a statistically significant increase in the vocabulary knowledge of low ability learners, following the use of Quizlet.

Hypothesis 2.1: Within the 4-day RI (earlier post-test), the shorter-spaced practice (1-day ISI=25% of RI) leads to higher post-test scores than the longer-spaced practice (7-day ISI=175% of RI), where the shorter spacing group completes the earlier delayed post-test at the optimal time.

Hypothesis 2.2: Within the 28-day RI (later post-test), the longer-spaced practice (7-day ISI=25% of RI) leads to higher scores than the shorter-spaced practice (1-day ISI=3.6% of RI) in the post-tests, where the longer spacing group completes the later delayed post-test at the optimal time.

Determining the research questions and hypotheses is considered as the starting point for any research. In turn, the selection of an applicable research methodology and methods will be directed by the research questions (Wellington, 2000). The following section justifies and discusses the research methodology adopted in this present study.

3.3 Rationale for the Selected Research Methodology

3.3.1 Research Framework

In Chapter Two, the relevant literature was reviewed regarding the importance of the lag effect for promoting vocabulary retention by comparing two different spacing schedules, namely longer and shorter practice distribution (see subsection 2.4.1). Also, examined in this area were previous studies on the benefits of technology integration, specifically using a variety of activities provided by Quizlet (see subsection 2.4.2.4). This review is intended to produce evidence upon which a theoretical framework can be constructed for the present study. The framework in question investigates the existing research on preferences for spacing distribution and on the integration of technology into the vocabulary learning process. In particular, an experimental method is implemented here, specifically a quasi-experimental design (see subsection 3.3.4) to enable a comparison of two spacing distributions for learning vocabulary and to test the research hypotheses. This type of comparison between methods or groups, together with hypothesis testing, is typical of experimental studies (Royse, 2004).

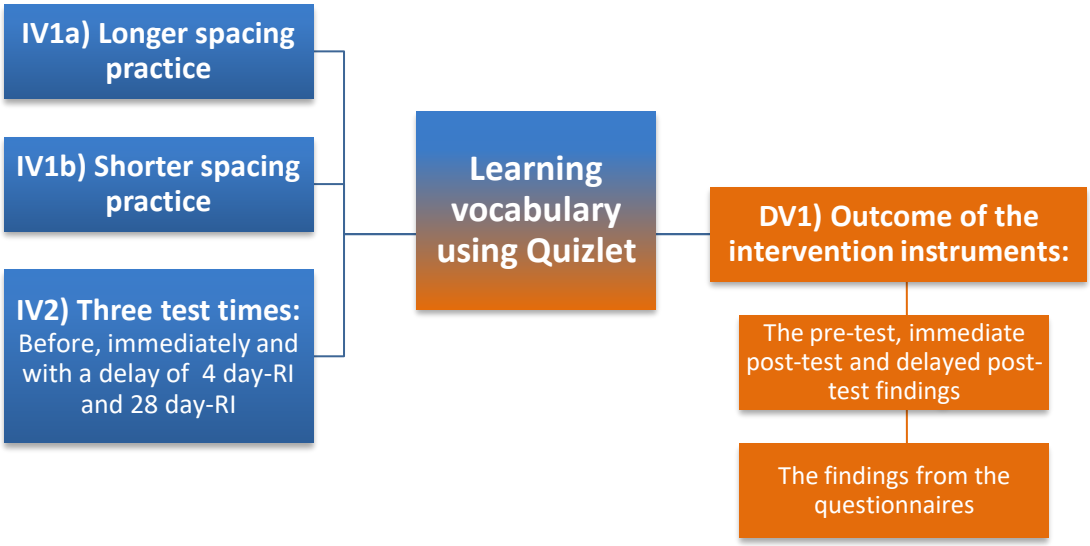
However, as stated by Louis et al. (2007), “the fundamental purpose of experimental design is to impose control over conditions that would otherwise cloud the true effects of the independent variables upon the dependent variables” (p.155). Johnson and Christensen (2008) describe independent variables as variables that a researcher will manipulate when conducting an experiment, in order to comprehend their possible influence on the dependent variable. Therefore, the development of the research framework in the present study required the independent and dependent variables to be precisely determined, as explained below.

Two independent variables were applied in this study. The first of these was a between-subjects factor. This means that two treatment conditions, comprising longer- and shorter-spaced sessions, were imposed upon the learners during the intervention. The second was a within-subjects factor, indicating the three test times at which the vocabulary tests were administered: directly before the start of the intervention, immediately after its completion, and in a delayed post-test.

Meanwhile, the dependent variables in this study refer to the outcomes of using the intervention instruments, namely a vocabulary test administered as a pre-test, immediate post-test, and delayed post-test, and a questionnaire developed to assess the perceived effectiveness of using Quizlet, completed after the intervention. A vocabulary level test (VLT), background questionnaire (completed as a pre-test) and working memory tests (WMTs, following the completion of all treatment sessions) were also used to gather further information about the sample and to confirm that they had a comparable level of ability and background. The pre-tests, and immediate and delayed post-tests were then administered to measure and compare the students' improvement in both groups (in the longer- and shorter-spaced sessions) to ascertain the optimal between-session spacing for teaching and learning vocabulary, with a view to improving students' vocabulary retention. Finally, the findings from the Quizlet questionnaires were utilised to discover the perceived effectiveness of using Quizlet for vocabulary learning in a typical classroom setting. Figure 3.1 reviews the potential relationship between the independent and dependent variables.

Figure 3.1

The Research Framework for the Independent Variables: between-subjects Factor (IV1a, IV1b), within-subjects Factor (IV2) and Dependent Variables (DV1)



3.3.2 Theoretical Underpinning

Constructivism is a learning theory that is particularly associated with psychology (Bada & Olusegun, 2015), with roots in philosophy as well as psychology. It is a learning theory that was first formulated by Dewey (1938) and then further developed by Piaget (1977), who introduced it into the study of child development. Later, Vygotsky (1978) added social and cultural impact to theories on the learning process and the role of these factors in the construction of knowledge (cited in Allsop, 2016). Constructivists believe that learners build an understanding of the surrounding world and find aids to help them comprehend their experiences (Norton & Wiburg, 2003). Conversely, teachers coach and guide learners to improve and evaluate their knowledge and understanding. Therefore, the key focus of constructivism is student-centred learning; i.e. learners build their own understanding and knowledge (Bada & Olusegun, 2015). However, three core epistemological categories of constructivism are proposed by Glasersfeld (1984, 1990): cognitive constructivism (where knowledge is actively constructed), radical constructivism (where all knowledge is constructed) and social constructivism (where knowledge is socially constructed) (Doolittle, 2007; McLeod, 2019).

The research design of this current study (see section 3.3.4) is consistent with the constructivist theoretical perspective, because the Quizlet activities were adopted with the

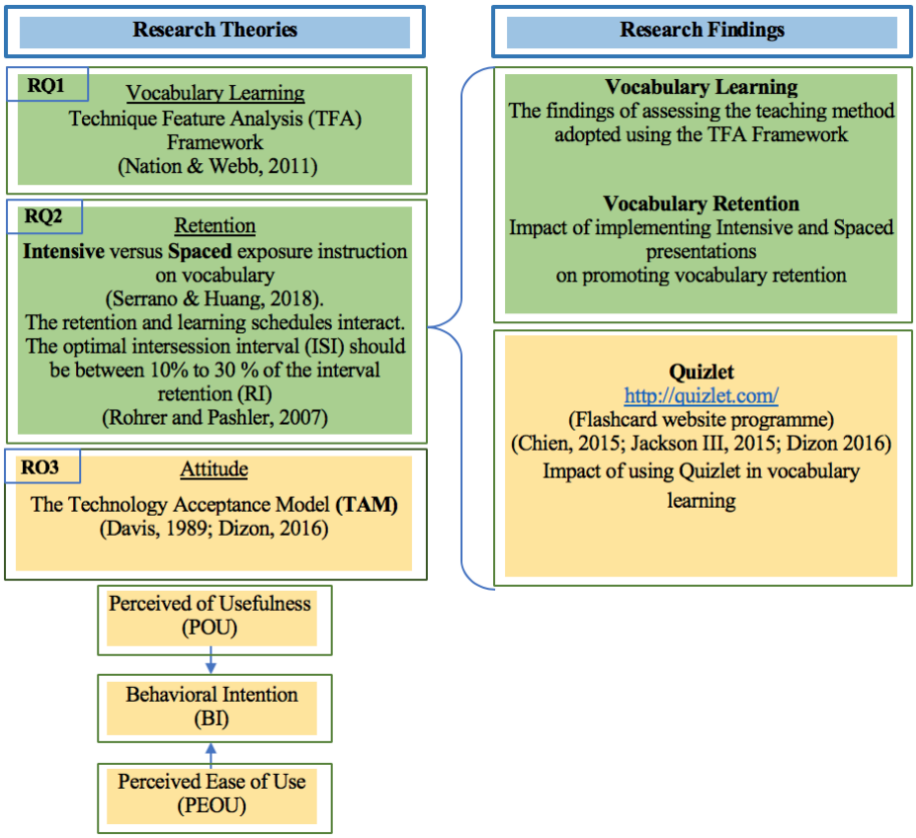
intention that they would actively engage the students in the learning process. Thus, they constructed their vocabulary knowledge through their engagement, which facilitated an increase in their vocabulary knowledge. Their engagement was confirmed by their scores in performance tests, which are explained in detail in Chapter Four. This engagement is corresponding to the epistemological category of cognitive constructivism. Additionally, the researcher acted as a facilitator to guide the learners' use of Quizlet. This role of the researcher is related to the second epistemological category of radical constructivism, in which the knowledge is actively built by learners. Therefore, constructivist learning theory formed the theoretical underpinning of this study, with regard to the teaching and learning of vocabulary using technology (Quizlet). The next subsection explains the theoretical and conceptual framework, followed by the research design of the current study.

3.3.3 Theoretical and Conceptual Framework

The researcher implemented several theories and frameworks to answer the three research questions, as well as justify and evaluate the quality of the teaching method used in the current research. Figure 3.2 provides a summary of the main theories used, and also highlights the general research findings, providing with some key studies in line with the current thesis. The following figure provides a full picture of the theoretical and conceptual framework of the present thesis.

Figure 3.2

Theoretical and Conceptual Framework



3.3.4 Research Design

According to Cohen et al. (2002), interpretivism and positivism are the most commonly adopted paradigms in the social science context. These are characterised as the two main approaches for gathering and analysing social science research data, whereby qualitative data are typically applied in an interpretivist paradigm; quantitative data in a positivist paradigm, and both qualitative and quantitative data in a mixed methods approach. As the social context cannot be fully understood by investigating it objectively from the outside, it must also be explored subjectively from within by studying individuals' direct experiences of a specific setting (Cohen et al., 2011). Therefore, the philosophical position adopted in this study is a mixed methods approach, combining interpretivism and the post-positivist paradigms. Punch (2009) holds a paradigm to be

> …a set of assumptions about the social world, and about what constitutes proper teaching and topics for inquiring into that world; a set of basic beliefs, a world-

> view, a view of how science should be done (ontology, epistemology, methodology). (p.358)

For any research to be valid, it is essential that it is underpinned by a suitable philosophical assumption or paradigm, in order to identify the appropriate research methods. Therefore, it is important that a researcher selects an appropriate paradigm to ensure that a study is successful and produces meaningful and useful findings. The view of reality adopted here, based on the educational research literature, can be seen to rest on two main assumptions: scientific (positivist) and interpretivist. Both are relevant to this study. From a positivist perspective, reality exists "out there, driven by immutable natural laws" (Guba, 1990, p.19), where "objects have an independent existence and are not dependent for it on the knower" (Cohen et al., 2007, p.7). The present researcher holds that the most appropriate epistemological perspective for this research was positivist, because it was important to assess the treatment of interest (i.e. longer versus shorter spacing distribution), delivered within a real practice setting (i.e. the classroom), in as consistent a manner as possible, given the limitations of a teaching environment. The choice of a positivist approach must also be accompanied by as consistent an assessment as possible, within the constraints of a classroom that lacks the control that is afforded by a laboratory environment. Because Quizlet has already been used worldwide, it is appropriate to conduct this study on its application in practice, thereby making a more interesting contribution to knowledge than it would have, if limited to a clinical context.

Conversely, interpretivism is an anti-positivist model, which relies on the notion that different individuals in society adopt diverse attitudes and influences, associated with their own experience. They build their ideas of a phenomenon based on their individual understanding of it (Hatch & Cunliffe, 2006). Therefore, interpretivism was selected as an epistemological stance in this study to investigate the participants' beliefs. Here, students are the focus of the learning process and teachers are the focus of practical knowledge. It is essential for educational researchers to consider teachers' views in the design and implementation of any courses. Moreover, an interpretive epistemology is mainly supported by a qualitative paradigm (Crotty, 2005). Therefore, open-ended questions and one-to-one interviews were used in this study, as the perceptions of both students and their teachers were highlighted.

The combination of research methods adopted in this study involved the use of both qualitative and quantitative research methods. The rationale behind choosing such a mixed methods approach was to answer the research questions as thoroughly and appropriately as possible, while taking advantage of various analytical tools (Lyons, 2015). This approach was also intended to validate the notion that

> …we can often learn more about our research topic if we can combine the strength of methods focused on quantitative data with the strength of methods focused on qualitative data, while compensating at the same time for the weaknesses of each method. (Punch & Oancea, 2014, p.339)

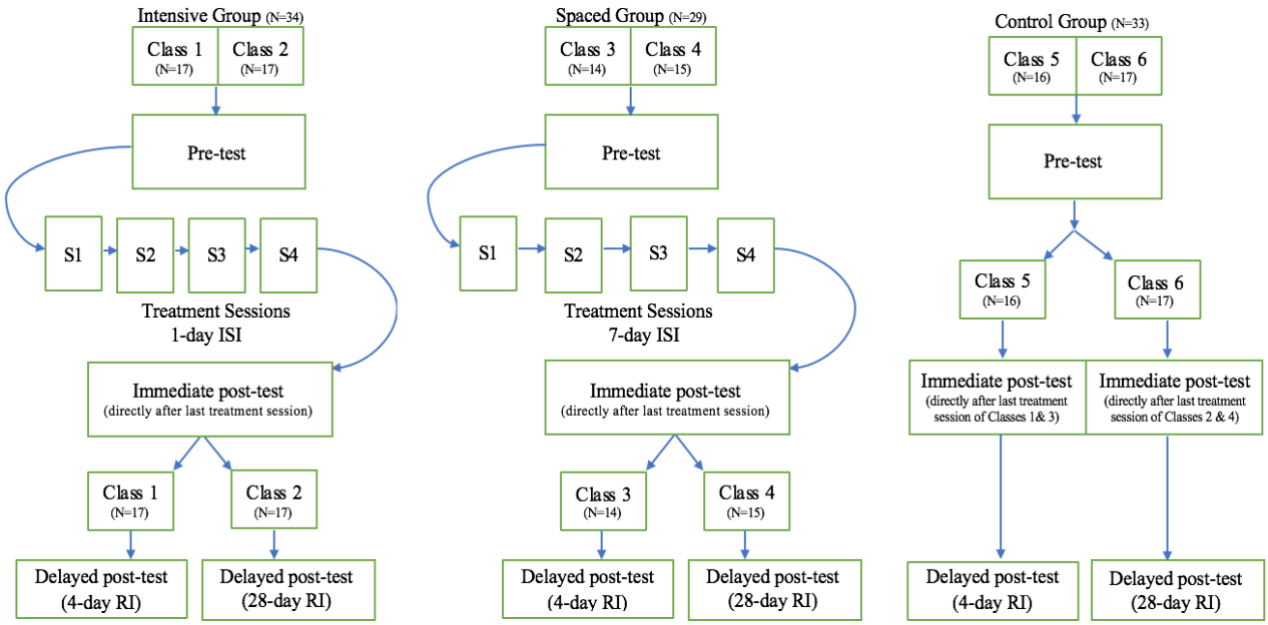
Pring (2015, p.80) supports this view, given that "One purpose of research is to explain what is the case or what has happened. A reason for seeking explanations might be to predict what will happen in the future or what would happen if there were to be certain interventions". Thus, the present researcher used a mixed methods, experimental design to measure and test the participants' performance in an intervention trial, so that an answer or explanation could be developed to start filling the knowledge gap identified. However, there are other aspects to consider; for example, it is important for a researcher to value the participants' subjective experiences, in order to gain a better understanding of the impact of implementing the treatment in question within a real-world, classroom setting.

To summarise, a mixed methods approach, combining two fundamental paradigms (positivist and interpretivist) was adopted in this study, after deciding upon the implementation of both qualitative and quantitative approaches (Creswell & Clark, 2017). Quantitative methods refer to a numerical or statistical approach in a research design, i.e. they represent quasi-experimental research; meanwhile, interpretivism supports the use of qualitative methods, such as one-to-one interviews. Therefore, these philosophical assumptions supported the researcher in answering the research questions, designing the research methodology, determining the data collection methods, and selecting the study sample. In this current research, the research questions and hypotheses have been addressed according to the philosophical framework (see section 3.2).

In considering the paradigmatic stance of this study and to address its research questions, it was identified that a combination of qualitative and quantitative methods was

required. This mixed methods approach was selected to enhance the strengths and minimise the weaknesses of each method (Creswell & Clark, 2017). Figure 3.3 illustrates the research design of the present thesis.

Figure 3.3
Research Design



3.4 Research Methodology

3.4.1 Research Context: College of Technology

With permission, the current study was conducted at the English Language Center (ELC) at a College of Technology, which is one of the primary providers of technological education under the administration and management of the Sultanate of Oman's Ministry of Manpower. The College is located in Oman. However, the teachers at the College come from different backgrounds and from all over the world.

Every year, the College receives around 800 students, who have graduated from Secondary schools at the age of 18 or 19. The number of students enrolling at the College changes from year to year, depending on the educational preferences and aims of these High School graduates in the region. Students on the ELC's Foundation Programme tend to share a similar educational background in EFL, and generally come from non-English speaking

backgrounds, whereby their exposure to the English language has mainly taken place in the classroom. However, they are also highly likely to encounter English via the internet and other media that are accessible to most of these students, both within and outside the College. Moreover, many will already speak another language apart from Arabic, such as a local language; for example, Jabbali (Shahria), Mahri or Habioot.

The Foundation Programme encompasses four levels: pre-elementary (low-beginner), elementary (beginner), intermediate, and advanced. It is organised around an academic year consisting of three semesters, with the first and second semester comprising 14 weeks in total, separated by a two-week break for the students. Students of all levels receive 18 hours per week of basic English language tuition, i.e. writing, reading, listening and speaking, as well as one weekly study skills and time management session. Meanwhile, Basic and Pure Mathematics are taken by Level Two and Level Four students for four hours a week, while Level Three Students have four hours of IT per week. It should be noted here that English is the medium of instruction for both Mathematics and IT. As all the College's specialisations are taught through the medium of English, the Foundation Programme is aimed at equipping students with the necessary academic language skills to grasp their subjects and follow their courses.

A new intake of students enters the College at the beginning of both Semester One (September-December) and Semester Two (January-April), and a placement test[1] for English language proficiency is administered to them all, so that they can be placed at the correct level, according to their language skills. However, Semester Three (Summer) is a nine-week term, with no new student intake. Most Level 1 students in Semester Three will have previously failed the Level 1 test or re-joined the students who suspended their studies in Semester Two. Level 1 students take 18 hours per week of English instruction, including Reading (four hours), Writing (five hours), Grammar (four hours), and Listening and Speaking (five hours), with a further hour for study skills. Here, the students go to language

1 The placement test is used to measure the English language level of each new intake of students at the College. It mainly tests students' English grammar and vocabulary and includes 100 questions (multiple-choice questions and gap-filling exercises), with a point being awarded for each correct answer. Students are assigned to a level based on these test results (0-20=Level 1; 21-50=Level 2; 51-70=Level 3; 70-90=Level 4, and over 90=post-foundation). The duration of the test is 90 minutes.

laboratories every week to apply what they have learned in each module, using the Moodle online platform, which provides sets of exercises for each Level 1 module.

3.4.2 Research Sample

The study population consists of around 400 Level 1 students (with pre-elementary English proficiency) on the Foundation Programme at the College of Technology. A total of approximately 143 potential participants were invited to take part in this study, including four Level 1 teachers at the English Language Center (ELC) and their 139 students (male and female), aged between 18 and 19 years. The study sample was not representative, but was rather an opportunistic sample (Thompson, 2012) of students who were likely to attend a typical class at the College, across a wide range of students. This means that no students were selected according to their representativeness of the general College student population or of all the students enrolling in the English language classes, which is generally the case for class groups.

The sample was derived from six intact classes out of 16 taught by 16 English teachers on the Programme. In each class, there were around 25 students, as these classes are usually arranged to accommodate 20-30 students. The six classes selected were divided into three groups (shorter-spaced [Intensive], longer-spaced [Spaced]) and the Control Group; ensuring that each group contained approximately 30 students, when considering the possibility of excluding some students from the study because of their periods of absence. The researcher acknowledged that randomising the sample selection within each class would be ideal for the experimental design. Although this could be considered as a limitation to the research design, the reality that the researcher had to work within was based on a real practice setting. Therefore, the two classes in each experimental group were taught separately, due to the difficulty of combining them in a single classroom, given their different timetables and the shortage of sufficiently large empty classrooms. However, having more than one class in each of the experimental groups reduced the possible influence of extraneous variables (for example, different levels of teaching engagement in each class) on the study outcomes. This helped counter any bias within the data. It is also worth mentioning here that the three groups were approximately matched in terms of age and educational background (see Appendix 15). In addition, the students were mainly male, as males made up most of the College population.

Furthermore, a background questionnaire, VLT, and WMTs were administered to confirm the equivalence of the six classes. Further, a Control Group was included to verify that any improvement observed in the post- and delayed post-test amongst the Intensive and Spaced groups was not due to a test effect. This group was a test-only group, i.e. they did not receive any additional treatment during the intervention period but were merely provided with their regular teaching (Mackey & Gass, 2005). The Control Group was useful to illustrate that the Experimental Groups were unlikely to have access outside the intervention to the target vocabulary either pre-intervention or during the intervention. This gave the researcher the confidence about whether the changes in the Experimental Groups' performances were because of the intervention or something else. The researcher acknowledges that there might be some other possible variables that were not tested, or could affect the Control Group's performance, because these were different classes from those in the Experimental Groups. For example, any difference in the Control Group's performance could possibly be due to extraneous factors particular to the classes that were part of the Control Group, rather than because they did not receive the intervention. However, the researcher tried to minimise the impact of this possible limitation by implementing the background and baseline tests. The researcher also ensured having two classes in each group and sufficient sample size.

Finally, participation in this study was completely voluntary for both the students and the teachers. The teachers were selected purely according to their willingness to participate. However, the activities designed for this intervention formed part of their standard teaching provision for the students. Aside from this, the students had the right to withdraw from the research (but not the teaching activities) without repercussions and at any time during the project.

3.4.3 Data Collection

Different data collection instruments (qualitative and quantitative methods) were used to best answer the research questions. The main methods employed were interviews, questionnaires and performance tests (pre-, immediate and delayed), in addition to the flashcard software programme, Quizlet, which was implemented as a learning method (see Table 3.1).

Table 3.1

Summary of the Methods and Dates of the Data Collection, as well as the Process of Data Analysis

Research Questions	**Data Collection Methods/ Groups**	**Data Collection Date**	**Data Analysis**
1) Is Quizlet (a Computer-assisted Language Learning [CALL] tool) an effective programme for promoting vocabulary learning amongst low ability learners in the classroom?	Baseline tests prior to the intervention: - Background questionnaires, - VLT - WM tests - Pre-tests - Post-tests following the intervention: An immediate post-test and delayed post-test were administered to compare mean differences between the Intensive, Spaced and Control Groups.	September-October 2018 October-November 2018	**SPSS** - Descriptive statistics - Parametric: One-way ANOVA Two-way mixed ANOVA (*Post hoc* pairwise comparisons with Bonferroni) - Non-parametric: Kruskal-Wallis tests
2) Does the time distribution of the practice sessions (intensive versus spaced) moderate the benefits of using Quizlet to promote vocabulary learning and retention amongst low ability learners?	- Pre-tests - Post-tests following the intervention: The immediate post-test and delayed post-test can be submitted to compare mean differences between the two Experimental Groups (Intensive & Spaced).	November 2018	**SPSS** - Descriptive statistics - Parametric: a Two-way mixed ANOVA (*Post hoc* pairwise comparisons with Bonferroni) - Non-parametric: Kruskal-Wallis test Friedman's Two-Way ANOVA
3) What are the perceptions of teachers and students regarding the use and implementation of Quizlet?	- Semi-structured interviews (four core classroom teachers) - Questionnaires (Experimental Groups)	December 2018	**Thematic analysis** (qualitative data analysis) **SPSS** - Descriptive statistics - Parametric: An independent samples t-test - Non-parametric: The Mann-Whitney test

The first hypothesis (see section 3.2) predicts a significant positive impact on the vocabulary learning of low ability learners following the use of Quizlet. Thus, the researcher compared the mean differences between the two Experimental Groups (Intensive and Spaced) and the Control Group. Two hypotheses were linked with the second research question: the first predicting that the shorter-spaced practice (Intensive – one day between sessions) would assist vocabulary learning in the short term to a greater extent than the longer-spaced practice. Meanwhile, the second hypothesis predicted that longer-spaced practice (Spaced – seven days between sessions) would facilitate long-term retention. These hypotheses were based on previous literature on the lag effect (for example, Cepeda et al., 2006). The researcher compared the mean differences between the Intensive and Spaced Groups at each time point. Finally, to obtain an answer to the third research question, the researcher conducted semi-structured interviews with the four core classroom teachers who had volunteered to participate in this study and administered questionnaires (which included open-ended questions) to the students, in order to gather deep information about their views of using Quizlet. The next section will review these tools in more detail.

3.4.4 Data Collection Tools

Several researchers have emphasised the importance of providing clear justification for the research methods and data collection tools selected in a study, in order to maintain the credibility of the research approach (Ritchie et al., 2013). The following sections will therefore highlight and justify the data collection tools implemented here in more detail.

3.4.4.1 Interviews

The first data collection method to be described in this subsection is the semi-structured interview (see Appendix 9). This method was selected to gain a rich understanding and deep information on the teachers' verbalised views and perceptions of using Quizlet for vocabulary learning. The researcher was the one who conducted the interviews with the teachers. The interview questions were based on the responses to the student questionnaire. Semi-structured interviews are flexible and include both open and closed-ended questions. Therefore, they are neither completely structured, nor designed to elicit yes/no answers. However, neither do they give the interviewer full freedom to lead the direction of the discussion, but rather to probe responses of interest and elicit clarification (Bryman, 2015).

It was planned to interview the teachers individually for approximately 30 minutes each at the College and at a time that was suitable for them. After obtaining ethical consent, these interviews were audio-recorded and transcribed with the permission of the interviewees. To conduct an effective interview, it was important to establish a relationship with the interviewees by making a series of visits to the case study institution beforehand (Marshal & Rossman, 2011). The present researcher had worked as a lecturer at this institution in the past and so she was familiar with it and had a good relationship with the teachers there. However, direct interaction between a researcher and participant in an interview can lead to misunderstandings and feelings of anxiety (Kvale, 2008). Therefore, it was decided to add three open-ended questions to the questionnaire instead of interviewing the students, in order to gather deep information from all the participants. These open-ended questions enquired about ease of use (PEOU), the usefulness of use (PU) Quizlet, and the students' behavioural intention (BI) to use it in future.

3.4.4.2 Questionnaires

Background and Quizlet questionnaires were distributed in this study to gain further information about the participants (see Appendix 7). The background questionnaire was implemented before the treatment to gather background information about the students (i.e. gender, age, languages) and data on their language use (i.e. daily, weekly, etc.) (Freed et al., 2004; see also Appendix 15 for more details of the findings on the participants' background information). The Quizlet questionnaire was then deployed after the intervention in an immediate post-test to investigate the participants' perceptions of using Quizlet, specifically with regard to its usefulness and ease of use, and how this affected their intentional behaviour to use it in future. The questionnaire was to be adapted from a questionnaire used by Dizon (2016), and it was conducted to answer the third research question. For this purpose, it was translated into Arabic for the students (see Appendix 8). It is worth noting here that the students were able to give responses in their L1 (Arabic), but the translated data are presented in the Results Chapter (see section 5.2).

The Quizlet questionnaire was divided into three main themes: perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI), according to which the results will be organised. Each theme included 8-9 items, with one open-ended questionnaire question, as shown in Table 3.2.

Table 3.2
The Quizlet Questionnaire Themes

Theme	No. of Items	Survey Items	Open-ended Questionnaire Question
PEOU	8	6, 9, 12, 15, 16, 18, 21, 25	2nd question
PU	9	1, 4, 8, 11, 14, 17, 20, 23, 24	1st question
BI	8	2, 3, 5, 7, 10, 13, 19, 22	3rd question

Note: For an explanation of each survey item and open-ended questionnaire question, see Appendix 7

The Quizlet questionnaire, as shown in Table 3.2, included both quantitative (survey) and qualitative (open-ended questionnaire question) data. In total, the survey comprised 25 items (statements), distributed across the three identified themes. The responses to these items were rated using a five-point Likert scale (Strongly disagree=1, Disagree=2, Not sure=3, Agree=4, and Strongly agree=5). It should also be mentioned here that the survey contained a number of negative items, for which the order of scoring was reversed, so that high values were maintained for positive responses to all items. Józsa and Morgan (2017, p.9) mention that

> When computing the summated scale, the negatively worded items need to be reversed in terms of the weighting. For example, in case of a 5-point scale strongly disagree is given a weight of 5 and strongly agree is given a weight of 1.

Moreover, the quantitative questionnaire data were analysed using IBM SPSS Statistics software (Version 24). Meanwhile, qualitative questionnaire data were generated using three open-ended questionnaire questions; the responses to these were analysed manually by organising them into themes (Braun & Clarke, 2006), whereby each question was related to one of three primary themes.

3.4.4.3 The Intervention

This study was quasi-experimental in nature and participation in that it involved an English vocabulary-learning activity, using a software flashcard programme called Quizlet with Level

1 English language learners. These learners were taught by the researcher and observed by their core classroom teachers on four occasions, over a period of four days for the Intensive Group and over four weeks for the Spaced Group. In particular, they used a Quizlet study set, created by the researcher to learn 34 new noun words (see subsection 3.4.4.9). These target words were conveyed to the students through several activities provided by Quizlet, accessed via the following link: https://quizlet.com/_40rkad (see subsection 3.4.4.8 for a description of these activities). The above-mentioned vocabulary learning activities were to be undertaken during regular lessons in the ELC language labs.

The intervention described above lasted for four weeks at a rate of one session per week for 20 minutes in the Spaced Group, and one session per day for 20 minutes, on four consecutive days for the Intensive Group. The entire class from each Experimental Group were sitting in a single language laboratory, each with their own PC and headsets, and they were learning at the same time. The lab technician was also available to help in solving any technical problems with PCs or the internet that could happen during the intervention, which lasted 20 minutes in any session. Generally, the intervention went smoothly with the help of the lab technician and the core teacher being given to the researcher to ensure that each student was able to do the activities of each session without fail and with the same amount of time for each class. The environment was generally quiet, except for the researcher giving instructions or helping a student. Both the Experimental Groups and the Control Group were then administered immediate and delayed post-tests. After the intervention, the students were asked to complete a questionnaire on the perceived effectiveness of Quizlet (see subsection 3.4.4.2) and the teachers were invited to attend semi-structured interviews (see subsection 3.4.4.1); see Appendix 20 for more detail about the logistic plan for the intervention of the current study.

3.4.4.4 Tests

Before undertaking the treatment conditions, the participants took a VLT and completed a background questionnaire. The 2,000-5,000 VLT (Schmitt et al., 2001) assessed the participants' vocabulary size to enable the researcher to identify the participants' vocabulary knowledge. There were 30 questions in each VLT, with one mark awarded for every correct answer. Thus, the total maximum score for each individual VLT was 30 (see subsection 4.4.1 and Appendix 13). After the treatment conditions, the participants were also administered a

Working Memory Test (WMT) to assess their ability to memorise words (see subsection 4.4.2 and Appendix 11). For this purpose, digit span tests (forward and backward) were considered appropriate in this study (Climie & Rostad, 2011; Elsayyad, 2014). These baseline tests were aimed at ascertaining the homogeneity of the sample. In addition, the pre-, immediate post-, and delayed post-tests tested the participants' knowledge of 34 nouns (see subsection 3.4.4.9 and Appendix 12). These items were then divided into two main categories: Recognition (17 items) and Recall (17 items). 'Recognition' consisted of two main sections: Passive Recognition (eight multiple-choice questions from English into Arabic) and Active Recognition (nine multiple-choice questions from Arabic into English). 'Recall' also comprised two key sections: Passive Recall (eight translation of items from English into Arabic) and Active Recall (translation of nine items from Arabic into English) (see Appendix 10).

3.4.4.5 Pre-tests

The pre-test was a paper and pencil test (see Appendix 10), designed to test both productive and receptive knowledge (Laufer & Goldstein, 2004) of 34 noun words, taken from the Level 2 Vocabulary Log. The test consisted of four sections: receptive recall, productive recall, receptive recognition and productive recognition (Laufer & Goldstein, 2004). Each of the 'active recall' and 'active recognition' sections included nine items, and the 'passive recall' and 'passive recognition' sections contained eight items each. A paper-based test method was deemed appropriate here to avoid any amendments being made to the answers, as it is easy to obtain feedback and change answers in computer-based tests. The test was conducted to compare the findings for the three groups (Intensive, Spaced and Control) with other findings for the same groups in the latter two tests (the immediate and delayed post-tests), in order to determine their progress at different test points. The pre-test was used to establish that the participants had the same baseline knowledge prior to the intervention.

3.4.4.6 Immediate Post-test

The immediate post-test was similar to the pre-test, with the researcher keeping the type of question and formatting consistent but varying the order of the items within each section (see Appendix 10). The immediate post-test was administered to all groups directly after the fourth and final Quizlet activity. This test was used to compare the findings for the three groups (Intensive, Spaced and Control), in order to find out which group demonstrated better

vocabulary retention in the immediate post-test.

3.4.4.7 Delayed Post-test

The delayed post-test was also similar to the pre- and immediate post-tests, but with the items in a different order within each section. It was designed to measure the longer-term retention of vocabulary items, thus comparing the impact of the three different conditions (Intensive, Spaced and Control) on vocabulary recall. This test was administered four days after the final practice session to half the participants in each group, while the other half took the same test at four weeks (28-days), following the treatment conditions being brought to a close. Thus, half the participants in the Spaced and Intensive Groups completed the delayed post-test in the optimal ISI:RI ratio (25%), i.e. for the 1-day Group (Intensive): 4-day RI=25% and 28-day RI=3.6%, whereas for the 7-day Group (Spaced): 4-day RI=175% and 28-day RI=25%.

3.4.4.8 Use of a Software Programme: Quizlet

For this study, the researcher adapted Quizlet (http://quizlet.com/), a flashcard software programme that offers a range of vocabulary tasks, such as Flashcards, the Speller, Learn, Test, the Scatter Game and the Gravity Game (Quizlet, 2018). There were several reasons for choosing this particular programme. First, Quizlet is widely used, with over 140 million study sets and over 20 million active users every month across 130 countries. It is also compatible with both the Apple iOS and Google Android mobile platforms. Moreover, it is free to access and easy to use. In addition, Quizlet is mainly focused on vocabulary, which is the focus of this study. However, the main point in its favour is that it can be manipulated to suit individual contexts. This enables teachers to create their own study sets along with an audio-visual exposition, which is available in the website to facilitate vocabulary learning in terms of form and meaning.

In addition, Quizlet can either be used on a PC or mobile device. However, not all these activities or features are available in the Quizlet mobile application, as shown in Table 3.3 (Dizon, 2016).

Table 3.3
Features of the Quizlet Website and Mobile App

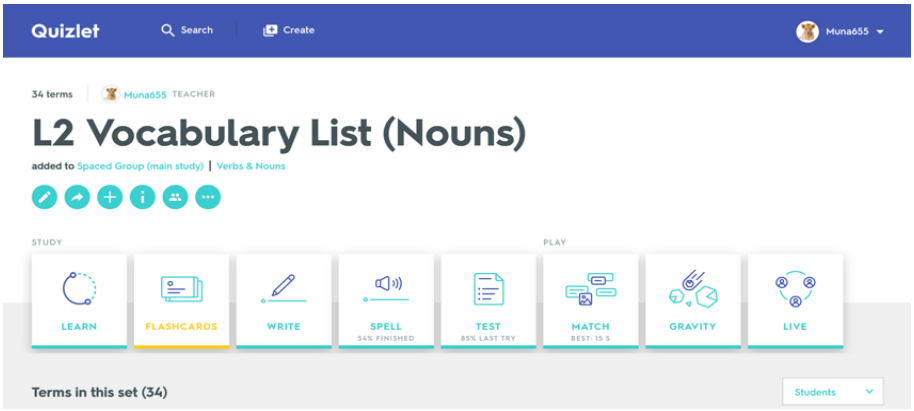
Feature	Website	Mobile app
Word lists	+	+
Flashcards	+	+
Speller	+	
Learn	+	+
Test	+	
Scatter game	+	+
Gravity game	+	

Furthermore, mobile app users receive fewer records of their progress and performance than those using the website (Dizon, 2016). Therefore, the students in this study accessed the programme on PCs at College. It was considered preferable to use the website, in order to benefit from these additional features, such as being able to track the students' progress, which enabled the students to be monitored by the researcher.

As mentioned earlier, the Quizlet website provides several vocabulary activities in the form of word lists and flashcards (Quizlet, 2018). It was essential to ensure that all the materials were balanced and that both Experimental Groups (Intensive and Spaced) had the same amount of exposure to the target vocabulary items in each training session and across all the sessions. Therefore, the researcher administered the following training material, which was designed to control the total number of repetitions of words, the types of activities (i.e. recognition and recall), the total amount of time spent, and the number and range of vocabulary tasks (i.e. Flashcards, Match, Spell, Write & Test), within the four training sessions. Figures 3.5, 3.6, 3.7 & 3.8 explain the training material that was designed for learning vocabulary in each of the four sessions attended by both groups (Intensive and Spaced).

Figure 3.4. shows the browser-based version of the flashcard set landing page for 34 noun words taken from Level 2 (L2) Log Vocabulary (see subsection 3.4.4.9).

Figure 3.4
Quizlet's Flashcard Set Landing Page



First Session: In Session One, the students completed two tasks (Flashcards and Matching), mainly focused on recognition activities. This session lasted for 20 minutes. (see Table 3.4 and Figures 3.5 & 3.6).

Table 3.4
Design of Session One

	Task/Mode	Learning Method	Type	Time Spent	No. of Repetitions
1.	Flashcards	- Automatic-play - Individual	Recognition	3 mins	1
2.	Flashcards	- Manual-play - In pairs	Recall/ Recognition	10 mins	2
3.	Match	- Individual	Recognition	7 mins	1
	Total			20 mins	4

Figure 3.5
Screenshot of Study Flashcards on Quizlet

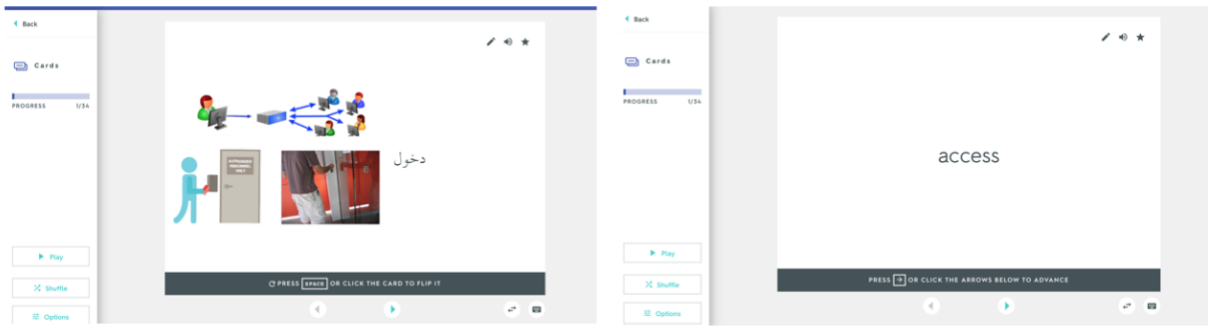


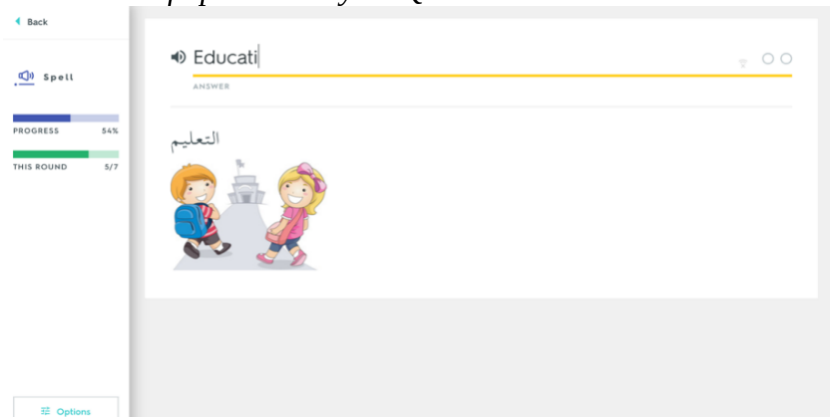
Figure 3.6
Screenshot of Match Activity on Quizlet

Second Session: In Session Two, the students began by revising the study set and then completed the Spelling activity. The activities included recognition and recall, as illustrated in Table 3.5 and Figure 3.7.

Table 3.5
Design of Session Two

	Task/Mode	Learning Method	Type	Time Spent	No. of Repetitions
1.	Flashcards	- Automatic-play - Individual	Recognition	3 mins	1
2.	Spell	- Spelling the words in English - Audio: On & Slow	Recall/ Recognition	17 mins	2
	Total			20 mins	3

Figure 3.7
Screenshot of Spell Activity on Quizlet

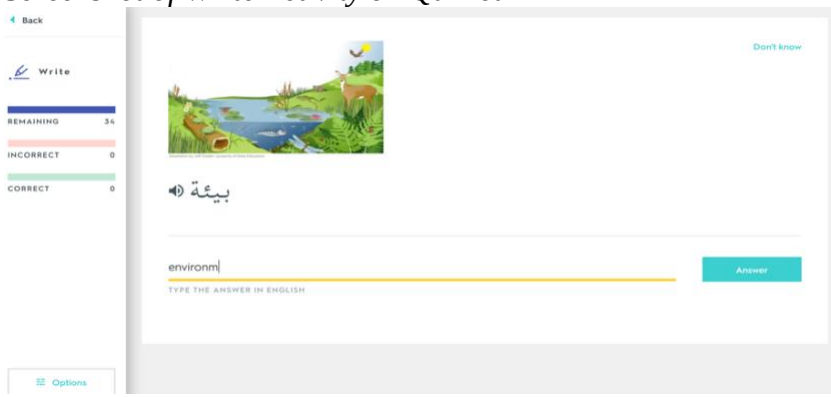


Third Session: In Session Three, the students revised the study set and practised writing with Write. Again, the activities included recognition and recall, as shown in Table 3.6 and Figure 3.8.

Table 3.6
Design of Session Three

Task/Mode	Learning Method	Type	Time Spent	No. of Repetitions
1. Flashcards	- Automatic-play Individual	Recognition	3 mins	1
2. Write	- Writing word forms in English - Audio: On	Recall	17 mins	2
Total			20 mins	3

Figure 3.8
Screenshot of Write Activity on Quizlet



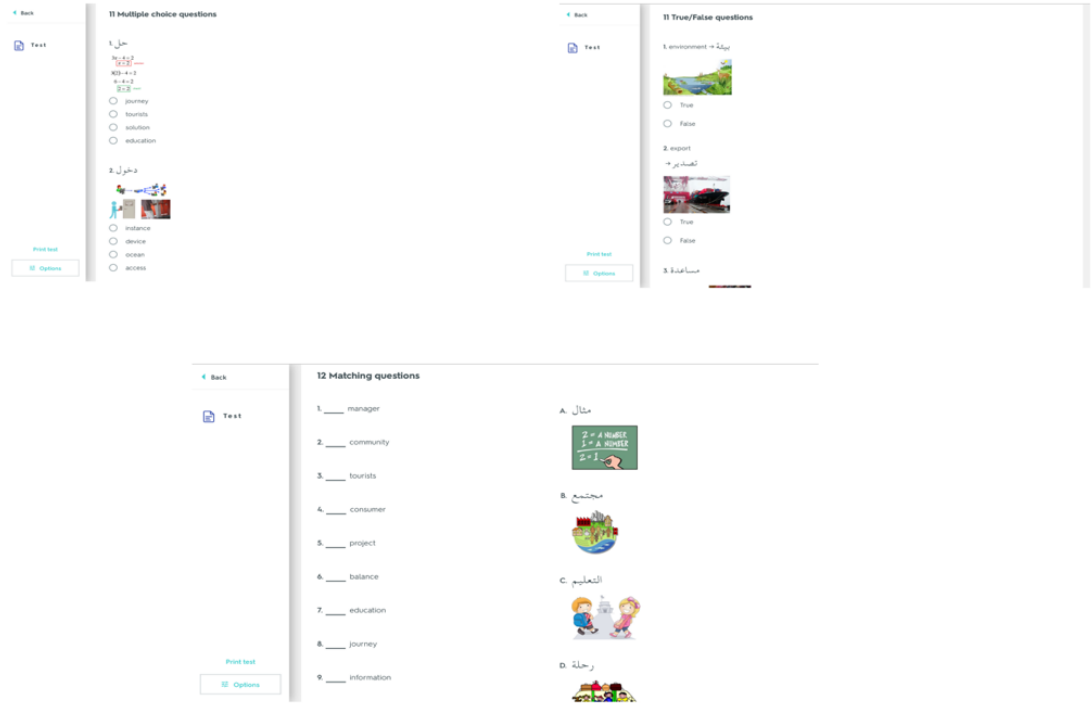
Fourth Session: In Session Four, the students revised Flashcards and then completed the Test task, focusing on recognition and recall knowledge (see Table 3.7 and Figure 3.9).

Table 3.7
Design of Session Four

Task/Mode	Learning Method	Type	Time Spent	No. of Repetitions
1. Flashcards	- Automatic-play Individual	Recognition	3 mins	1
2. Test	Multiple choice, matching, writing, true/false questions	Recall/ Recognition	17 mins	2
Total			20 mins	3

Figure 3.9

Screenshot of Test on Quizlet



In order to evaluate the Quizlet vocabulary learning activities used in this present study, the researcher demonstrated the features of the Technique Feature Analysis (TFA) Framework, as formulated by Nation and Webb (2011) (see Table 3.8). Hu and Nassaji (2016) recommend using the TFA Framework to evaluate the effectiveness of any vocabulary learning activity. This framework is considered to be a successful predictive indicator of the depth of processing facilitated by the activity (see subsection 2.3.5). Hu and Nassaji advise the inclusion of more features of TFA in vocabulary learning activities. Therefore, the use of the TFA Framework was essential for assessing the teaching method adopted in this present study, using Quizlet in practice (see Table 3.8).

Table 3.8
Five-component Framework of Assessment Criteria (Adopted from Nation & Webb, 2011, p.7)

Criteria	Scores	%
Motivation		
Is there a clear vocabulary learning goal?	1	
Does the activity motivate learning?	1	66.7
Do the learners select the words?	0	
Noticing		
Does the activity focus attention on the target words?	1	
Does the activity raise awareness of new vocabulary learning?	1	100
Does the activity involve retrieval negotiation?	1	
Retrieval		
Does the activity involve retrieval of the word?	1	
Is it productive retrieval?	1	
Is it recall?	1	100
Are there multiple retrievals of each word?	1	
Is there spacing between retrievals?	1	
Generation		
Does the activity involve generative use?	0	
Is it productive?	1	33.3
Is there marked change that involves the use of other words?	0	
Retention		
Does the activity ensure successful linking of form and meaning?	1	
Does the activity involve instantiation?	0	75
Does the activity involve imaging?	1	
Does the activity avoid interference?	1	
Total score	**14**	**77.8%**
Maximum score	18	100%

Table 3.8 shows that the teaching method adopted, namely the use of Quizlet and spacing between sessions, met 14 out of 18 criteria related to the main five components. This covers around 78% of the features of the TFA Framework. Therefore, this framework provides a rationale for adopting the current research method, where explicit vocabulary is learned through Quizlet.

3.4.4.9 Target Words

The researcher selected 34 target words (see Appendix 12) based on specific criteria. First, all the noun words belonged to the 2,000-word level and academic word list, and were selected from the Level 2 Vocabulary Log (see Appendix 14), i.e. a list of words taken from the Level 2 course book, in order to ensure that the target words were above the participants' current

level of vocabulary knowledge (i.e. Level 1). The Level 2 Vocabulary Log consists of 100 words, representing different parts of speech (see subsection 1.2.3 for further information about using the Vocabulary Log at the College). The total number of nouns falling into the stated range was 34, all of which were taken as the target words for the current study. It is worth mentioning that the researcher used the *Compleat Lexical Tutor* website (https://www.lextutor.ca/vp/comp/) to identify the frequency and range of vocabulary from an adaptation of Heatley et al.'s (2002) range.

A Vocabulary Level test (VLT) (Schmitt et al., 2001) was also administered to the participants to ensure that their level of vocabulary knowledge was categorised at below the identified vocabulary range for the intervention. It was also important to find equivalent meanings in the students' L1 (i.e. Arabic), so that the words were clearer for the students. Lastly, only one part of speech was selected for this study, because nouns constitute one of the main components of any sentence and all everyday speech (Webb, 2005).

In addition, studies on word learning argue that some parts of speech are harder to learn than others (see, for example, Childers & Tomasello, 2006). Therefore, the researcher chose the noun, which represented most of the words on the Level 2 list, in keeping with the specified criteria. This was to ensure that the target words were as similar as possible in their frequency and range, while also presenting a degree of challenge to the learners. It consequently avoided the effect of having widely differing levels of difficulty between target words.

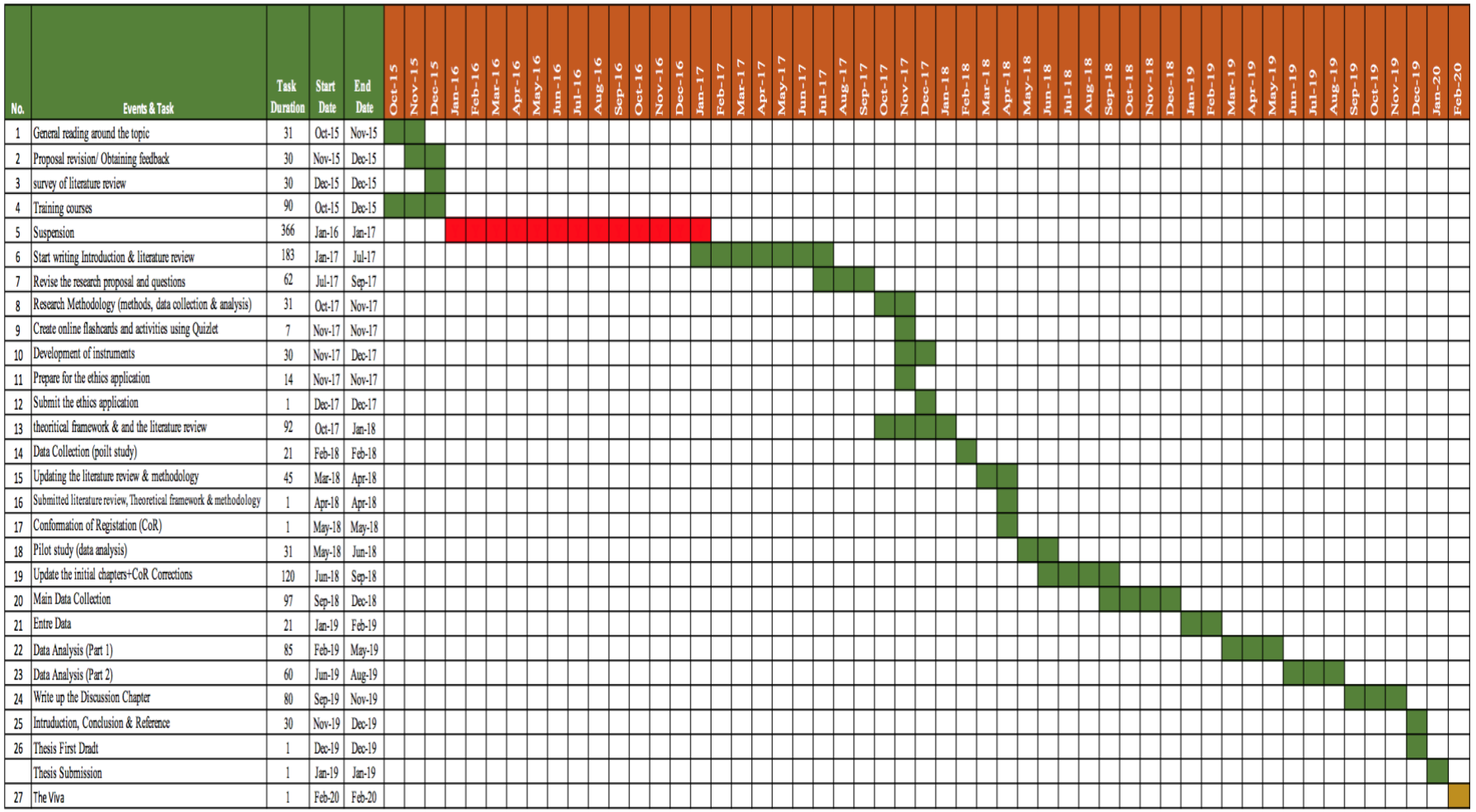
3.4.4.10 Planning and time scales

A Gantt Chart showing the planning and time scales of when each research task and event took place by each month is presented in Figure 3.10.

Title of study: Evaluating the Effectiveness of Spaced Practice using Computer Assisted Language Learning (CALL) in Teaching and Learning English Vocabulary in the Classroom: The Case of Oman

NAME	DURATION	TART DA	END DATE
Muna Muqalbal	4 YEARS	Oct-15	Feb-20

No.	Events & Task	Task Duration	Start Date	End Date
1	General reading around the topic	31	Oct-15	Nov-15
2	Proposal revision/ Obtaining feedback	30	Nov-15	Dec-15
3	survey of literature review	30	Dec-15	Dec-15
4	Training courses	90	Oct-15	Dec-15
5	Suspension	366	Jan-16	Jan-17
6	Start writing Introduction & literature review	183	Jan-17	Jul-17
7	Revise the research proposal and questions	62	Jul-17	Sep-17
8	Research Methodology (methods, data collection & analysis)	31	Oct-17	Nov-17
9	Create online flashcards and activities using Quizlet	7	Nov-17	Nov-17
10	Development of instruments	30	Nov-17	Dec-17
11	Prepare for the ethics application	14	Nov-17	Nov-17
12	Submit the ethics application	1	Dec-17	Dec-17
13	theoritical framework & and the literature review	92	Oct-17	Jan-18
14	Data Collection (poilt study)	21	Feb-18	Feb-18
15	Updating the literature review & methodology	45	Mar-18	Apr-18
16	Submitted literature review, Theoretical framework & methodology	1	Apr-18	Apr-18
17	Conformation of Registation (CoR)	1	May-18	May-18
18	Pilot study (data analysis)	31	May-18	Jun-18
19	Update the initial chapters+CoR Corrections	120	Jun-18	Sep-18
20	Main Data Collection	97	Sep-18	Dec-18
21	Entre Data	21	Jan-19	Feb-19
22	Data Analysis (Part 1)	85	Feb-19	May-19
23	Data Analysis (Part 2)	60	Jun-19	Aug-19
24	Write up the Discussion Chapter	80	Sep-19	Nov-19
25	Intruduction, Conclusion & Reference	30	Nov-19	Dec-19
26	Thesis First Dradt	1	Dec-19	Dec-19
	Thesis Submission	1	Jan-19	Jan-19
27	The Viva	1	Feb-20	Feb-20



COMPLETED
SUSPENSION
PENDING

Figure 3.10
The Research Gantt Chart

3.4.5 Pilot Study

A pilot study was conducted to test and evaluate the research instruments, in order to increase their validity. It was undertaken directly after receiving ethical approval (see Appendix 1 and section 3.7 for additional details). This process took three weeks during the Spring Term of 2018 (see Appendices 22 & 23). The number of participants in the pilot study initially comprised one teacher and 33 Level 1 students from the Foundation Programme. At the end of the treatment conditions, the number of participants was reduced to 16, because of students being absent from at least one study phase. Table 3.9 presents the planned timeline for the pilot study intervention.

Table 3.9
Planning for the Pilot Study Intervention

Level 1 students (N)	**Massed Group (8)**	**Spaced Group (8)**
Week 1	Baseline tests: Working Memory Test (Digit Span) Background Questionnaire Vocabulary Language Tests Pre-test	Baseline tests: Working Memory Test (Digit Span) Background Questionnaire Vocabulary Language Tests Pre-test **Intervention 1**
Week 2	**Intervention 1 & 2** Immediate post-test	**Intervention 2** Immediate post-test
Week 3	Delayed post-test Questionnaire	Delayed post-test Questionnaire

Both groups (Spaced and Massed) were comparable in number and ability, as they all had low English language ability and poor vocabulary knowledge, as per the pre-test. Table 3.10 gives a full description of the participants' performance in both groups at the three test points (pre-, post- and delayed post-tests).

Table 3.10
Descriptive Statistics for the Data at Each Time Test Point for Both Groups (Spaced and Massed)

Group		Pre-test		Post-test		Delayed Post-test	
		Recognition	**Recall**	**Recognition**	**Recall**	**Recognition**	**Recall**
Massed	Mean (*SD*)	3.88 (*4.82*)	/	8.63 (*1.60*)	2.13 (*2.48*)	7.88 (*1.36*)	1.13 (*0.84*)
	95% CI	[-0.16, 7.91]	/	[7.29, 9.96]	[0.06, 4.19]	[6.74, 9.01]	[0.43, 1.82]
	Min-Max	0.00-12.00	/	5.00-10.00	0.00-6.00	6.00-9.00	0.00-2.00
Spaced	Mean (*SD*)	4.50 (*2.14*)	0.88 (*0.84*)	9.38 (*0.92*)	3.25 (*2.49*)	9.13 (*1.64*)	3.25 (*2.25*)
	95% CI	[2.71, 6.29]	[0.18, 1.57]	[8.61, 10.14]	[1.17, 5.33]	[7.75, 10.05]	[1.37, 5.13]
	Min-Max	1.00-7.00	0.00-2.00	8.00-10.00	0.00-7.00	6.00-10.00	1.00-6.00

Note: The maximum test score was 20.00. / - The Massed Group did not give any correct answers for the pre-test recall.

Recognition output: Two-way mixed ANOVA was conducted to identify whether the two groups progressed differently between the three test time points. The results indicate a main effect of Time: $F(1.34, 28) = 23.34$, $p < .001$, $\eta_{p2} = .63$. However, the main effect of the Group was not statistically significant: $F(1, 14) = 1.11$, $p = .31$, $\eta_{p2} = .073$. Neither was the Group * Time interaction statistically significant: $F(1.34, 28) = .091$, $p = .84$, $\eta_{p2} = .006$.

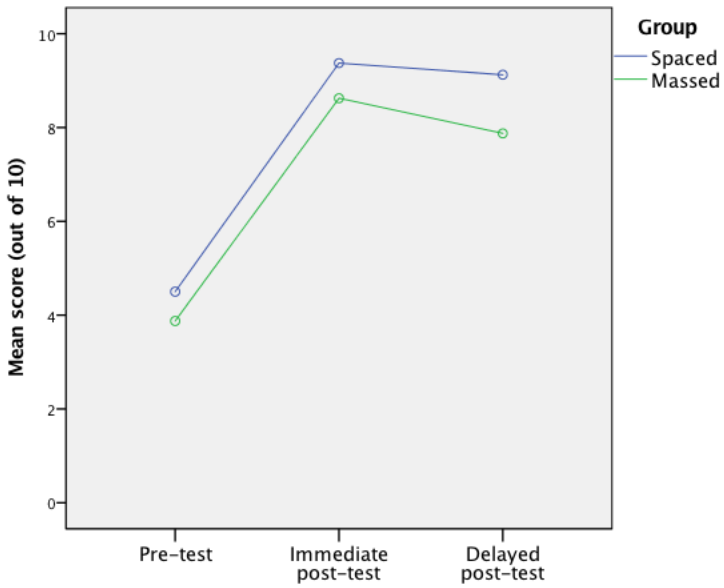
Post hoc pairwise comparisons with Bonferroni corrections were then conducted to explore whether there was any improvement between different test time points within each group. Both the Spaced and Massed groups made significant progress between Time 1 and Time 2 and between Time 1 and Time 3. However, there were no significant differences between Time 2 and 3 for either group (see Table 3.11 for more detail).

Table 3.11
Recognition Output (Two-way Mixed ANOVA)

Group		Time 2	Time 3
Spaced	Time 1	$p = .003$, $d = 2.96$	$p = .013$, $d = 2.43$
	Time 2		$p = 1.00$, $d = 0.00$
Massed	Time 1	$p = .003$, $d = 1.32$	$p = .032$, $d = 1.13$
	Time 2		$p = .77$, $d = -0.51$

The results in Figure 3.11 show that both conditions produced equally satisfactory results, with no differences in the short term. Conversely, the longer-term results favoured the Spaced Group, with 2.7% loss, while the Massed Group were found to have lost 8.7% of the target words.

Figure 3.11
Estimated Marginal Means of Recognition



Recall output: Results derived from two-way mixed ANOVA indicate that there was a main effect of Time: $F(1.44, 28) = 12.90$, $p < .001$, $\eta_{p2} = .48$. The main Group effect was, however, not statistically significant: $F(1, 14) = 3.82$, $p = .071$, $\eta_{p2} = .214$. Neither was the Group * Time interaction found to be statistically significant: $F(1.44, 28) = 1.01$, $p = .84$, $\eta_{p2} = .067$.

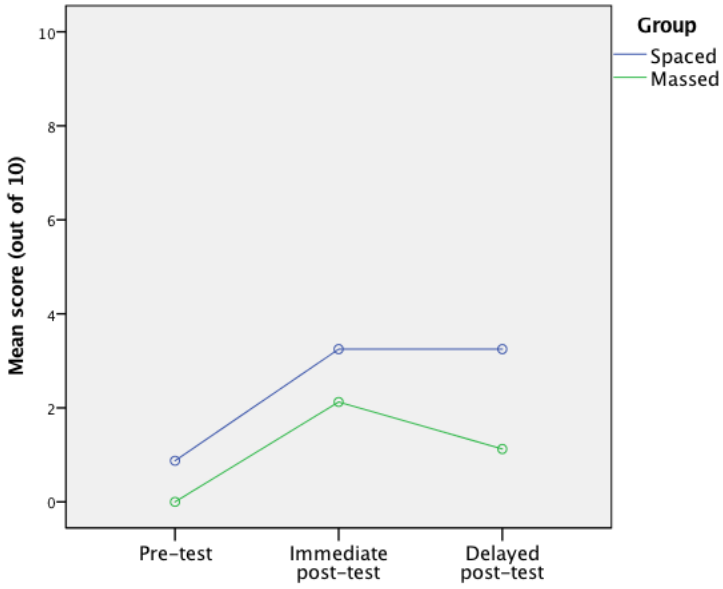
Post hoc pairwise comparisons with Bonferroni corrections showed that the Massed Group made no significant progress between Time 1 and Time 2, or between Time 1 and Time 3. However, the Spaced Group made significant progress between Time 1 and Time 2 and also between Time 1 and Time 3. However, there were no significant differences between Time 2 and 3 for either group, as explained in Table 3.12.

Table 3.12
Recall Output (Two-way Mixed ANOVA)

Group		Time 2	Time 3
Spaced	Time 1	$p < .040$, $d = 1.28$	$p < .002$, $d = 1.40$
	Time 2		$p = 1.00$, $d = -0.19$
Massed	Time 1	$p = .072$, $d = 1.22$	$p = .189$, $d = 1.90$
	Time 2		$p = .247$, $d = -0.54$

The results in Figure 3.12 reveal that both conditions improved significantly, with no differences in the short term. However, the longer-term results showed preference for the Spaced Group, in that they lost nothing, while the Massed Group were found to have lost 47% of the target words.

Figure 3.12
Estimated Marginal Means of Recall



The results showed no significant difference between the two Groups (Spaced and Massed) in their Recognition knowledge; they had both progressed by the immediate test and preserved their progress in the delayed post-test. However, in terms of Recall knowledge, the

Spaced Group performed far better than the Massed Group, who showed no significant changes over the three time points. Therefore, this result is in line with the previous literature, which supports spacing between training sessions (see subsection 2.4.1).

As shown in Table 3.13, both groups expressed a positive opinion of using Quizlet, in terms of its perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI) to use it in future.

Table 3.13
The Results of the Questionnaire for the Pilot Study

	Usefulness		**Ease of Use**		**Behavioural Intention**	
	Spaced	Massed	Spaced	Massed	Spaced	Massed
Median	40.50	26.00	30.50	28.00	30.50	22.00
Min-Max	38.00-43.00	23.00-36.00	28.00-34.00	19.00-31.00	27.00-35.00	18.00-29.00

Note: A participant could potentially yield the highest rating=50/lowest rating=10.

Looking at the ease of using Quizlet, the Spaced Group was more positive than the Massed Group: U=5.50, z= -1.96, p= .48. The Spaced Group was also more positive about their intention of using Quizlet in future: U=10.00, z= -2.09, p= .40. However, there was no difference between the two groups' attitudes to the usefulness of Quizlet: U=9.50, z= -1.92, p= .054.

The pilot study results showed that the Spaced Group generally performed better than the Massed Group in the delayed post-test. This finding is consistent with study-phase retrieval theory (Greene, 1989) and the findings of many existing studies (Kornell, 2009; Lotfolahi & Salehi, 2017; Nakata &Webb, 2016), which have revealed the advantages of spaced over massed practice for longer-term knowledge retention. In addition, both groups had positive perceptions of using Quizlet to learn vocabulary, which is in line with the findings of several researchers (Dizon, 2016; Jackson III, 2015). However, because of the short timeframe allocated to this pilot study, the researcher did not investigate optimal spacing between the sessions, as the aim of piloting at this stage was to assess the research tools. Therefore, while the pilot study compared Spaced versus Massed study sessions, the

main study would compare longer (7-day) versus shorter (1-day) spacing between practice sessions, so as to investigate the lag effect for vocabulary learning, as explained previously (see subsection 3.4.4.3).

Although there were several limitations to this pilot study, relating to the small sample size and the small number of lexical items used (see Appendix 21), it was significant for obtaining a full picture of the issues that were likely to face the researcher in the main data collection. It also provided a means of informing the research design and improving the research instruments. Therefore, the researcher increased the number of lexical items to 34 target words, which would help obtain more accurate results and reduce any side effects that could influence the findings. The pilot study also helped the researcher to set the intervention timeline and determine the time required for each activity.

3.5 Data Analysis

As mentioned in the research design section, the researcher in the present study collected both quantitative data (the VLTs, WMTs, performance tests and survey) and qualitative data (open-ended questionnaire items, one-to-one interviews), in order to answer the research questions (see section 3.2). For the purpose of analysing these data, IBM SPSS Statistics software (Version 24) was used to process the quantitative data, while thematic analysis was used for the qualitative data. Before starting to analyse the data, however, a marking scheme was devised, which is described in the following subsection.

3.5.1 Marking Scheme for Performance Tests

The first step, after gathering all the quantitative data, was to develop a marking scheme for the performance tests (pre-, immediate, and delayed), with a total number of 34 noun words per test. Each test was divided into four parts. The first two parts were Active Recall (9 words to be translated from Arabic into English) and Passive Recall (8 words to be translated from English into Arabic). The Passive and Active Recall parts were marked in two different ways. Firstly, the students were given no points (0) for an incorrect answer; 1 point for a correct but misspelled word; 2 points for a correct word, spelled correctly, and 99 points for no attempt. For example, the word, ‘Hunter’ could be marked as follows: Hunt=0, Hunte=1, Hunter=2, no attempt=99. This first marking scheme was applied when identifying the nature of the

students' attempted answers. Its purpose was to check the students' productive vocabulary knowledge as they attempted to produce the correct spelling (recalling the form) and correct meaning (recognising the meaning). Additionally, the students' engagement with the tests over the three time points was evaluated through consideration of missing answers and attempts to answer (see subsection 4.5.1). Secondly, for the analysis of the participants' overall performance over time, the researcher recoded the scores as incorrect = 0 (which included all responses initially coded as 0, 1, 99) and correct = 1 (including those answers initially coded as 2). Therefore, only two scores (1 and 0) were entered into ANOVA and other inferential statistics. The second marking scheme was adopted from the College, being the one that is normally used there for vocabulary spelling tests.

The other two parts of the performance tests (Active and Passive Recognition) included 17 multiple-choice questions (9 Arabic words, with 4 English definitions for each Arabic word) and English into Arabic tasks (8 English words with 4 Arabic definitions for each English word). The marking of these parts of the test was very straightforward, with one point awarded for every correct choice and zero awarded for an incorrect choice from the four available options.

Although the researcher scored the tests manually, an independent person (acting as the second marker) scored the tests in the same way. This second marker (a Linguistics PhD student who was also an Arabic native speaker) verified and ensured the reliability of the first marking, carried out by the researcher. This second marker checked 30% of the tests, following the same marking scheme. There were no differences between the researcher and the second marker in the way that they scored the tests, so there was 100% inter-rater reliability.

3.5.2 Quantitative Data Analysis

SPSS software was used to analyse the data generated via quantitative methods (the pre-tests, immediate and delayed post-tests, quantitative questionnaires). Initially, descriptive statistics (range, min-max scores, mode, mean or median, standard deviation) were employed for the quantitative data. These descriptive statistics were extracted to obtain a clearer idea of the performance of the three groups (Control, Intensive, Spaced). Descriptive statistics were also generated for the quantitative questionnaire data (survey) to obtain an overview of the

Experimental Groups' perceptions of Quizlet's effectiveness for vocabulary learning and retention.

Next, it was important for the researcher to check the normality of the data and the homogeneity of variance, as well as to detect any outliers. If the essential assumptions were satisfied, parametric statistical tests would be adopted, i.e. ANOVA and a t-test. Otherwise, non-parametric tests would be implemented, such as Kruskal-Wallis tests and Friedman's ANOVA, which would be conducted instead of ANOVA, and the Mann-Whitney test employed instead of the t-test. The checking of assumptions and justification of the selection of statistical tests for the quantitative data are provided in detail in the following section and Chapter Four.

There were two phases involved in data analysis. Before the intervention, a one-way ANOVA was implemented for the baseline tests (VLTs and WMTs), comparing the Intensive, Spaced and Control Groups. After the intervention, the data collected from the pre-test, immediate post-test and delayed post-test were submitted to Repeated Measures ANOVA, with *post hoc* tests to compare the mean differences between groups and within each group. This allowed the researcher to assess whether there were any significant differences (by calculating the *p* values) within or between the groups. However, it is insufficient to rely on *p* value only to interpret possible effects within the data; therefore, it is also important to calculate effect sizes to enable interpretation of the magnitude of change due to the intervention and between groups (Plonsky & Oswald, 2014). In fact, sample size (small or large) can affect the result of the *p* value, but has less influence on effect size (Wei et al., 2019). Therefore, the researcher used Plonsky and Oswald's (2014) benchmarks for *r* (*r*= .25 as small, *r*= .40 as medium, and *r*= .60 as large), and for Cohen's *d* for a between-subjects comparison (*d*= .40 as small, *d*= .70 as medium, . *d*= 1.00 as large) and a within-subjects comparison (*d*= .60 as small, *d*= 1.06 as medium, . *d*= 1.40 as large). Plonsky and Oswald's benchmarks are highly appropriate for L2 experimental-based research, as a means of interpreting effect size results (Wei et al., 2019). These benchmarks in the current study were applied to within-subject comparisons (comparing time points for the same learner) and between-subject comparisons (comparing two groups at one time point). To assess the reliability of effect sizes, 95% confidence intervals were calculated, in which the confidence intervals that do not cross zero were judged to be reliable indicators of an effect within the

data (Plonsky & Oswald, 2014). In addition, both partial and full eta-squared were reported for the ANOVA results (Norouzian & Plonsky, 2017).

3.5.3 Checking Assumptions

The design of the current study satisfies the requirement of a two-way mixed ANOVA, which was selected to explore any differences or change in performance among the independent variables, *Group* (three main experimental groups) and *Time* (three time-points), and to discover any two-way interaction between these two variables. Other assumptions that needed to be fulfilled before two-way mixed ANOVA could be run included homogeneity of variance, approximate normal distribution, and the absence of significant outliers (Field, 2018). Additionally, sphericity needed to be present, as it "measures whether differences between the variances of a single participant's data are equal" (Larson-Hall, 2010, p.336).

However, it is not enough to meet the requirement related to the research design; it is also important to consider other assumptions, mentioned earlier, that had to be met for a two-way mixed ANOVA. Therefore, the assumption of normality of distribution and homogeneity of variance was checked using the Shapiro-Wilk test and Levene's tests ($p > .05$) for the sub-sets of data that were analysed (i.e. Recall/Recognition, separately), with respect to all the 4-day and 28-day RI data gathered. By sub-setting the data as either 'Recognition' or 'Recall', the assumptions of normality and homogeneity were not completely fulfilled, as shown in Table 3.14.

Table 3.14
Shapiro-Wilk Test of Normality and Homogeneity of Variance for Recognition and Recall (4-day RI Sub-groups)

Treatment		Shapiro-Wilk Test		Levene's Test
		df	*P*-value	*P*-value
Pre-Recognition	Control	16	.746	.499
	Intensive	17	.565	
	Spaced	14	.216	
Pre-Recall	Control	16	.001	.269
	Intensive	17	.001	
	Spaced	14	.001	
Immediate Post-Recognition	Control	16	.503	.002
	Intensive	17	.001	
	Spaced	14	.001	
Immediate Post-Recall	Control	16	.001	.014
	Intensive	17	.262	
	Spaced	14	.843	
Delayed Post-Recognition	Control	16	.824	.003
	Intensive	17	.001	
	Spaced	14	.002	
Delayed Post-Recall	Control	16	.001	.131
	Intensive	17	.660	
	Spaced	14	.944	

Table 3.14 shows that the Control Group satisfied the assumption of normality for Recognition in the pre-, immediate post-, and delayed post-tests. However, this assumption was violated in all Recall sections of the three tests taken by the Control Group. In contrast, the results achieved by the Intensive and Spaced Groups were normally distributed for pre-test Recognition, immediate post-test Recall, and delayed post-test Recall. Conversely, the assumption of normality was violated by the Intensive and Spaced Groups in their pre-test Recall, immediate post-test Recognition, and delayed post-test Recognition. Similarly, the results of homogeneity of variance showed that the three treatments (pre-test Recognition, pre-test Recall, and delayed post-test Recall) satisfied the assumption, whereas the other three (immediate post-test Recognition, immediate post-test Recall, and delayed post-test Recognition) violated the assumption of homogeneity of variance. This meant that only half the results for the 4-day RI sub-groups satisfied the assumptions. The assumption of sphericity was also checked using Mauchly's test: $\chi_2(2) = 88.88$, $p < .001$. The assumption of sphericity for the 4-day RI sub-groups was violated, but according to the Greenhouse-Geisser

correction, ε̂= .537. The Greenhouse-Geisser estimate was applied to render the violation more robust, as the result was less than 0.75 (Field, 2018).

Next, it was also necessary to assess the data from the 28-day RI sub-groups to check for normal distribution, homogeneity of variance, and sphericity (see Table 3.15). When the 28-day RI sub-group data were divided into Recognition and Recall for each main test, the assumptions of normality and homogeneity were satisfied in three of these tests and violated in another three. Moreover, the assumption of sphericity was violated in the 28-day RI sub-group data: (χ^2(2) = 53.80, p< .001), with a Greenhouse-Geisser correction of less than 0.75.

Table 3.15
Shapiro-Wilk Test of Normality and Homogeneity of Variance for Recognition and Recall (28-day RI Sub-groups)

Treatment		Shapiro-Wilk Test		Levene's Test
		df	*P*-value	*P*-value
Pre-test Recognition	Control	17	.910	.584
	Intensive	17	.548	
	Spaced	15	.426	
Pre-test Recall	Control	17	.001	.661
	Intensive	17	.001	
	Spaced	15	.001	
Immediate Post-test Recognition	Control	17	.682	.008
	Intensive	17	.001	
	Spaced	15	.001	
Immediate Post-test Recall	Control	17	.001	.002
	Intensive	17	.055	
	Spaced	15	.353	
Delayed Post-test Recognition	Control	17	.842	.194
	Intensive	17	.001	
	Spaced	15	.001	
Delayed Post-test Recall	Control	17	.001	.002
	Intensive	17	.069	
	Spaced	15	.496	

Ultimately, when Recognition and Recall were run separately in each test (the pre-, immediate post- and delayed post-tests) for the 4-day and 28-day RI sub-groups, the assumptions of normality and homogeneity of variance were partially violated. The assumption of sphericity was corrected using three corrections, based on estimates of sphericity advocated by Greenhouse and Geisser (1959). In addition, the design of the

intervention in this current study satisfied the first requirement, meaning that two-way mixed ANOVA could be run. This mixed ANOVA test was necessary for answering the second research question, especially as no non-parametric equivalent of two-way mixed ANOVA (Larson-Hall, 2010) could be used to test comparisons between three groups over different time points (Field, 2018).

Recognition and Recall were subsequently divided into Passive and Active, in order to evaluate the effect of the distribution of practice on the students' passive and active knowledge at the three time points. First, the Shapiro-Wilk test and Levene's test were applied to the data generated in the pre-test for the passive and active knowledge findings for all the 4-day RI and 28-day RI sub-groups (Control, Intensive, Spaced). Next, it was essential to check the normal distribution, homogeneity of variance and sphericity for the data of the 28-day RI sub-groups (see Table 3.16). The 28-day RI sub-group data were also divided into Passive/Active Recognition and Passive/Active Recall for each main test, to assess the assumptions and then analyse the data.

Table 3.16
Shapiro-Wilk Test of Normality and Homogeneity of Variance for Pre-tests

Sub-test	**Group**	**4-day RI**			**28-day RI**		
		Shapiro-Wilk Test		**Levene's Test**	**Shapiro-Wilk Test**		**Levene's Test**
		df	*P*-value	*P*-value	df	*P*-value	*P*-value
Pre-passive Recognition	Control	16	.078	.226	17	.462	.030
	Intensive	17	.165		17	.004	
	Spaced	14	.336		15	.021	
Pre-active Recognition	Control	16	.320	.971	17	.082	.432
	Intensive	17	.062		17	.400	
	Spaced	14	.371		15	.317	
Pre-passive Recall	Control	16	.001	.847	17	.001	.777
	Intensive	17	.001		17	.001	
	Spaced	14	.001		15	.001	
Pre-active Recall	Control	16	.001	.002	17	.001	.017
	Intensive	17	.001		17	.001	
	Spaced	14	.001		15	.001	

As shown in Table 3.16, the results indicate that the Pre-passive Recognition satisfied the assumptions of normality and homogeneity of variance (p>.05) for the 4-day RI sub-groups, but these assumptions were violated for the 28-day RI sub-groups (p<.05). For Pre-active Recognition, the assumptions were met for both sub-groups (4-day RI and 28-day RI). However, the assumption of normality was violated at Pre-passive and Pre-active Recall for both sub-groups. Moreover, both sub-groups satisfied the assumption of homogeneity of variance in Pre-passive Recall, but this assumption was violated in Pre-active Recall for both sub-groups.

Table 3.17
Shapiro-Wilk Test of Normality and Homogeneity of Variance for Immediate Post-tests

Sub-test	**Group**	**4-day RI**			**28-day RI**		
		Shapiro-Wilk Test		**Levene's Test**	**Shapiro-Wilk Test**		**Levene's Test**
		df	*P*-value	*P*-value	df	*P*-value	*P*-value
Immediate-passive Recognition	Control	16	.035	.166	17	.063	.019
	Intensive	17	.001		17	.001	
	Spaced	14	.001		15	.001	
Immediate-active Recognition	Control	16	.134	.001	17	.542	.007
	Intensive	17	.001		17	.001	
	Spaced	14	.001		15	.001	
Immediate-passive Recall	Control	16	.001	.001	17	.001	.001
	Intensive	17	.017		17	.010	
	Spaced	14	.252		15	.029	
Immediate-active Recall	Control	16	.001	.014	17	.001	.063
	Intensive	17	.683		17	.030	
	Spaced	14	.700		15	.328	

Table 3.17 shows that the assumption of normality in the immediate post-test for the 4-day RI and 28-day RI sub-groups was generally found to be violated in most of the sub-tests. Moreover, the homogeneity of variance was largely violated for both sub-groups. It was only satisfied in the immediate-passive Recognition for the 4-day RI sub-groups and in the immediate-active Recall for the 28-day RI sub-groups.

Table 3.18
Shapiro-Wilk Test of Normality and Homogeneity of Variance for Delayed Post-tests

Sub-test	Group	4-day RI			28-day RI		
		Shapiro-Wilk Test		Levene's Test	Shapiro-Wilk Test		Levene's Test
		df	*P*-value	*P*-value	df	*P*-value	*P*-value
Delayed-passive Recognition	Control	16	.076	.468	17	.168	.042
	Intensive	17	.001		17	.001	
	Spaced	14	.001		15	.001	
Delayed-active Recognition	Control	16	.897	.026	17	.536	.055
	Intensive	17	.001		17	.001	
	Spaced	14	.001		15	.001	
Delayed-passive Recall	Control	16	.001	.001	17	.001	.001
	Intensive	17	.142		17	.042	
	Spaced	14	.044		15	.097	
Delayed--active Recall	Control	16	.001	.002	17	.001	.002
	Intensive	17	.469		17	.136	
	Spaced	14	.816		15	.255	

Similarly, in the delayed post-test, the assumptions of normal distribution and homogeneity of variance for the 4-day RI and 28-day RI sub-groups were violated in most of the sub-tests (see Table 3.18). Furthermore, the assumption of sphericity in all sub-tests for the 4-day RI sub-groups was violated. Therefore, the Greenhouse-Geisser estimate was employed as an appropriate correction for almost all the tests that were found to be less than 0.75 (Field, 2018). Only Passive Recognition in the pre-, immediate post- and delayed post-test was found to be greater than 0.75. Therefore, Huynh-Feldt is the appropriate correction to restore robustness to the violation (Field, 2018). Consequently, non-parametric tests were implemented to check the results generated by the two-ANOVA analyses. The results of the non-parametric tests mirrored the findings of the two-way ANOVA.

3.5.4 Qualitative Data Analysis

The qualitative data were composed of the four semi-structured interviews with four core teachers and the three open-ended questionnaire questions administered to the students. According to Ritchie et al. (2013), "there are no clearly agreed rules or procedures for analysing qualitative data, but many different possible approaches" (p.270). Therefore, thematic analysis was employed here. This refers to "the method that works both to reflect

reality and to unpick or unravel the surface of 'reality'" (Braun & Clarke, 2006, p.81). Therefore, codes were assigned to the interview responses provided by the core teachers, and the students' responses to the three open-ended questions in the Quizlet questionnaire. These responses, shared by the teachers, were then arranged under common themes, using diagrams and tables (Huberman & Miles, 2002) to further clarify the essential data themes.

3.6 Reliability and Validity

Reliability and validity are crucial to social research, in order to demonstrate and assess the quality of a study (Teddlie & Tashakkori, 2009). However, these two concepts are more applicable to quantitative research, while trustworthiness is appropriate for qualitative research (Hammersley, 2007), reflecting its credibility, transferability, dependability and confirmability (Lincolin & Guba, 1985).

To improve the quality of this current study, a number of methods (questionnaires, semi-structured interviews, tests) were included in the research design to help enhance reliability and trustworthiness (Berg, 2016). Therefore, methodological triangulation was adopted in this research, with data being obtained from multiple sources (teachers, students) (see subsection 3.2.5). The use of triangulation and multiple measures would ensure the credibility of this research (Cohen et al., 2007). In addition, the researcher decided to use Cronbach's alpha to calculate the reliability of the tests (see section 4.3) and questionnaire. "Cronbach's alpha is a model of internal consistency reliability based on the average inter-item correlation of an instrument" (Rovai et al., 2014, p.545). The resulting Cronbach's alpha value for the questionnaire was found to be .729, as shown in Table 3.19. Generally, an alpha value of .7 or higher is considered acceptable as an indication of reliability in questionnaires, according to Rovai et al. (2012).

Table 3.19
Reliability Statistics for the Questionnaire

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	No. of Items
.729	.753	25

3.7 Ethical Issues

Ethical considerations are essential for social science research. Researchers therefore need to be aware that "ethical concerns should be at the forefront of any research project and should continue through the write-up and dissemination stages" (Wellington, 2000, p.3). The current research was conducted on young adult learners. Therefore, ethical considerations were given high priority by the researcher to protect the participants' rights and interests (Yin, 2014). The researcher was also responsible for preserving the participants from any mental or physical risks associated with taking part in the study (Bryman, 2015). Therefore, a number of ethical issues were taken into account in this study. First, the researcher obtained ethical approval from the Ethics Committee of the Institute of Education at the University of Reading (see Appendix 1). This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request. The researcher also obtained permission to conduct the research at the College from both the College Dean and the Head of its English Language Center (see Appendices 2 and 3).

In addition, all the participants (teachers, students) were given an Information Sheet, which provided them with the research focus, objectives, research methods and a description of their roles in this study (Denscombe, 2014) (see Appendices 4 and 5). The students were provided with an Arabic version (see Appendix 6) to ensure that all the information given to them was easy for them to understand. Initially, the researcher asked them to sign a Consent Form. All participants were informed that they could withdraw from the research without repercussions and at any time during the project.

All the data collected were subsequently held in the strictest confidence by the researcher (Cohen et al., 2011). Pseudonyms were also used to protect the participants' privacy (Berg, 2016). Furthermore, the researcher securely saved all electronic data on her own PC, using a non-shared password. All documentation was also kept in inaccessible cabinets, in a locked office. On completing the research, the researcher will destroy all participant data and records.

3.8 Summary

The purpose of this chapter was to identify the research paradigms, the theoretical underpinning and the theoretical and conceptual framework of the current study. This chapter has also provided explanations and a rationale for the research methods adopted, and planned the data collection procedures. The research analysis outlines the use of SPSS to analyse the findings of the quantitative research methods, such as questionnaires and tests. It also highlights the use of thematic methods of analysing the interview data collected from teachers in a qualitative research approach, and the data from the students' responses to open-ended questions.

The pilot study was significant for obtaining a full picture of the issues that were likely to face the researcher in the main data collection; it provided a means of informing the research design and indicating any changes that the researcher might need to make in response to feedback received from the students, with regard to the methods applied in the pilot study. Finally, ethical considerations were specified in this chapter, while subsequent chapters will concentrate on the data analysis, discussion of the findings, and conclusion to the study.

CHAPTER FOUR: RESULTS 1, EXPERIMENTAL LEARNING GAINS

4.1 Introduction

In the Methodology chapter, the data collection methods and procedures for implementing the study were explained (see section 3.4.3). This chapter will now examine the experimental data, applying a statistical analysis of the participants' performance in the pre-, immediate post- and delayed post-tests. First, it will outline the tests conducted before, during and after the intervention, as shown in Table 4.1. The research questions will also be restated, before summarising the results of tests for normality and homogeneity. Secondly, the baseline findings will be presented, followed by the experimental findings for the first and second research questions. Meanwhile, the participants' perceptions (from the questionnaires and interviews), aimed at addressing the two remaining research questions, will be presented in Chapter Five. The present chapter, however, will conclude with a summary of the experimental findings.

Table 4.1
Summary of Tests Conducted in the Experimental Study

Tests	**Groups**		
	Intensive	**Spaced**	**Control**
2,000 Vocabulary Language Test (VLT)	√	√	√
3,000 Vocabulary Language Test (VLT)	√	√	√
5,000 Vocabulary Language Test (VLT)	√	√	√
Working Memory Test (Forward)	√	√	√
Working Memory Test (Backward)	√	√	√
Pre-test	√	√	√
Immediate post-test	√	√	√
Delayed post-test	√	√	√

4.2 Restating the Research Questions

The Research Questions are as follows:

1. Is Quizlet (a Computer-assisted Language Learning [CALL] tool) an effective programme for promoting vocabulary learning amongst low ability learners in the classroom?
2. Does the time distribution of the practice sessions (intensive versus spaced) moderate the benefits of using Quizlet to promote vocabulary learning and retention in low ability learners?
3. What are the perceptions of teachers and students regarding the use and implementation of Quizlet?

4.3 Reliability of the Tests

The reliability of the quantitative instruments was tested before beginning any exploration of the data. Cronbach's alpha (Cronbach, 1951) was used to check the reliability of the questionnaire, as described in Chapter Three (see section 3.6), as well as to calculate the level of internal consistency across 34 test items at the three time points (pre-, immediate post- and delayed post-test). The results of the pre-test indicated a reliable internal consistency of .84. Similarly, both the immediate and delayed post-tests demonstrated high reliability of .96. (see Table 4.2).

Table 4.2
Reliability Statistics Derived from the Performance Tests

Tests	**Cronbach's alpha**	**No. of Items**
Pre-test	.84	*34
Post-test	.96	*34
Delayed post-test	.96	*34

*The 34 noun words used to test the students at the three time points (see section 3.4.4.9).

It is worth noting here that the researcher is aware of other researchers (for example, McNeish, 2018) who have explored alternatives to Cronbach's alpha as a means of assessing

reliability. However, the reasons for selecting Cronbach's alpha in this study were based on the fact that it is the most frequently used method of measuring the reliability of scales and tests (Taber, 2018). The results for reliability, obtained using Cronbach's alpha in this present study, showed high internal consistency.

4.4 The Participants and Baseline Tests

As mentioned previously, the present study design is quasi-experimental, which is common in the field of educational research, especially as it results in low interference in existing school programmes (Porte, 2002). The six intact classes selected were assigned to three different conditions, i.e. Intensive (34 students), Spaced (29 students) and Control (33 students). These groups were assigned by the students' teachers, who had volunteered to participate in the study, and the college administration, based on the teachers' timetables and availability of computer laboratories (see section 3.4.2).

The initial number of participants agreeing to participate in this intervention was 139 students, who signed the consent form and took the baseline tests. However, the objective of this study was to apply an intervention in a real-world setting. Thus, it was recognised that in practice, student absences are to be expected. As such, it was decided that participants would be excluded from the quantitative analysis of the experimental learning gains if they missed one or more intervention sessions and/or one or more of the testing sessions. This decision was based on two different premises: first, the pre-test and immediate post-test were conducted on the same day, as the first and final (fourth) practice session, respectively. Therefore, in missing one of these tests, a student would also be missing one practice session. As a result, the data would be incomplete for these students, meaning that the data for answering the first and second research questions would not be accurate. In addition, missing a session would mean that the student concerned would not have received the same training as the other participants, who had attended all of the practice sessions. This could subsequently affect the learning gains of those students with missing sessions. Therefore, all the participants attended every practice session, with 96 students taking the tests. The data from these tests (pre-, immediate post- and delayed post-test) were then used to address the first and second research questions.

The three treatment groups were required to take three (2,000, 3,000 and 5,000)

vocabulary level tests (VLTs), and two working memory (WM) tests (forward and backward), as baseline tests for the intervention (see section 3.4.4.4). The VLTs were used to check whether the participants displayed the same level of vocabulary knowledge across the three groups. Meanwhile, the WM tests were employed to find out if there was any variation in the students' ability to recall information. Therefore, the following data analysis starts with an exploration of the descriptive statistics, followed by the results of the baseline tests (VLTs and WM tests).

4.4.1 Vocabulary Level Tests (VLT)

The first baseline test administered was version two of the 2,000-5,000 VLT (Schmitt et al., 2001). Descriptive statistics were generated for the three groups, in order to be able to process the results of the baseline test. As clarified earlier in the description of the research design, each Experimental Group consisted of two sub-groups of participants; the first sub-group completed the delayed post-test after a 4-day retrieval interval (RI) and the second sub-group completed the delayed post-test after a 28-day RI. Therefore, all test results in the present study were analysed separately for each sub-group and will be presented separately in the following sections.

4.4.1.1 The 4-day Retrieval Interval (RI) sub-groups

The descriptive statistics for the 4-day RI sub-groups (within the Control, Intensive and Spaced Groups), revealed that the participants in the Control Group scored 1.00, 0.00 and 0.00 median scores, the Intensive Group recorded median scores of 2.00, 0.00 and 0.00, and the Spaced Group achieved median scores of 2.50, 1.00, .00 and 2.5, respectively, derived from the 2,000, 3,000 and 5,000 VLTs. However, it was found that there was only one student who could get above 50% of the total scores in the 2,000 VLT, and this participant was in the Control Group, getting a score of 17 out of 30. Excluding this result, the descriptive statistics indicate that all the students participating in the study demonstrated a similar vocabulary level, indicating that the size of vocabulary across the groups corresponded to a word frequency level of less than 2,000. Their mean scores and other descriptive statistics are presented in Table 4.3.

Table 4.3
Descriptive Statistics for the VLTs (4-day RI Groups)

Group (N)		2,000 VLT	3,000 VLT	5,000 VLT
Control (16)	Mean (*SD*)	2.00 (*4.21*)	0.88 (*3.24*)	0.19 (*.75*)
	Median	1.00	0.00	0.00
	Min-Max	0.00, 17.00	0.00, 13.00	0.00, 3.00
	Range	17.00	13.00	3.00
Intensive (17)	Mean (*SD*)	2.59 (*2.12*)	1.12 (*1.41*)	0.53 (*1.38*)
	Median	2.00	0.00	0.00
	Min-Max	0.00, 7.00	0.00, 4.00	0.00, 5.00
	Range	7.00	4.00	5.00
Spaced (14)	Mean (*SD*)	2.86 (*2.18*)	1.43 (*1.60*)	0.64 (*1.50*)
	Median	2.50	1.00	0.00
	Min-Max	0.00, 8.00	0.00, 5.00	0.00, 5.00
	Range	8.00	5.00	5.00

Note: Maximum score for each VLT is 30.00.

One-way ANOVA was identified as the ideal test to check for any significant differences in the VLT scores between the groups. To run this test, three assumptions must be fulfilled: approximate normal distribution, no significant outliers, and homogeneity of variance (Field, 2018). For all the VLTs, the assumption of homogeneity of variance was found to be satisfied, using Levene's test ($p > .05$). However, in general, the results of the Shapiro-Wilk test ($p < .01$) were not found to be normally distributed across the groups, with the boxplots displaying outliers in the Control Group for all the VLTs (see Figures 4.1 & 4.2) and some outliers in the Intensive and Spaced groups in the 5,000 VLT (see Figure 4.3). Therefore, a Kruskal-Wallis test was employed instead of ANOVA to compare the VLT scores across the groups, due to this violation of normal distribution.

Figure 4.1
Boxplot of Test Scores on the 2,000 VLT for the 4-day RI Sub-groups

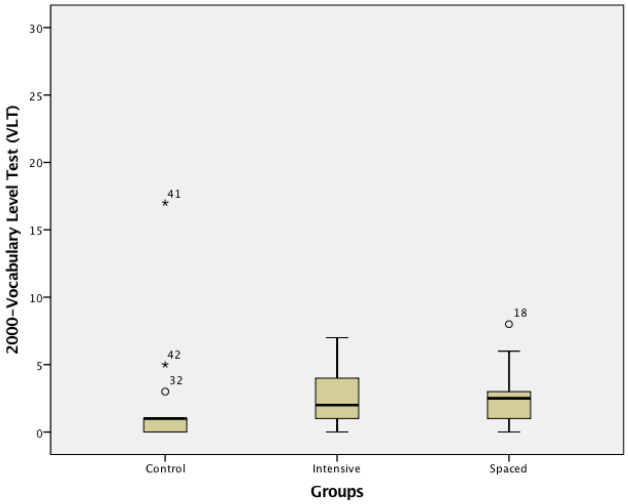


Figure 4.2
Boxplot of Test Scores on the 3,000 VLT for the 4-day RI Sub-groups

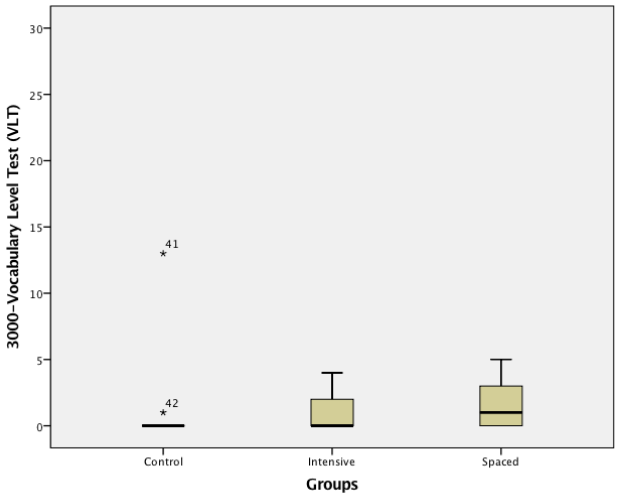
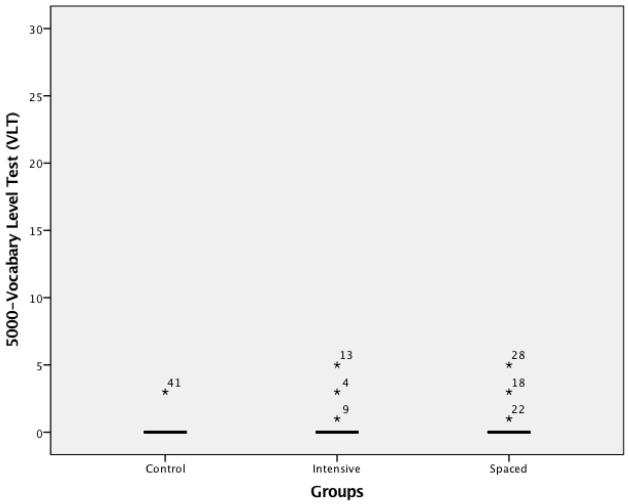


Figure 4.3
Boxplot of Test Scores on the 5,000 VLT for the 4-day RI Sub-groups



The Kruskal-Wallis test demonstrated significant differences between the participants' results for both the 2,000 and 3,000 VLTs: $H(2)=6.30$, $p= .043$. However, there were no significant differences found between the 5,000 VLT result scores: $H(2)=1.473$, $p=.479$. Therefore, pairwise comparisons with adjusted *p*-values were run for the groups to check for any differences between the Control vs. Intensive ($p= .150$, $r= -.34$), Control vs. Spaced ($p= .060$, $r= -.42$), and Intensive vs. Spaced ($p=1.000$, $r= -.08$) scores for the 2,000 VLT. The same comparisons were subsequently made between the Control vs. Intensive ($p= .173$, $r=-.33$), Control vs. Spaced ($p=.055$, $r=-.43$), and Intensive vs. Spaced ($p=1.000$, $r=-.10$) for the 3,000 VLT. The results revealed no statistical differences between the groups for the 2,000 and 3,000 VLTs, with small effect sizes. Finally, all of the 4-day RI sub-groups, all the groups displayed a similar level of vocabulary knowledge, which was below the 2,000-word frequency level. In fact, there was only one student (in the Control Group) who passed the 2,000 VLT. Nevertheless, the researcher decided not to delete the outliers, as the intention in this research was to conduct a study in a real-life setting, with the same differences appearing between the students as would usually occur in the classroom.

4.4.1.2 The 28-day Retrieval Interval (RI) sub-groups

In addition, the results of the VLTs for the 28-day RI sub-groups (within the Control, Intensive and Spaced groups) were analysed, starting with descriptive statistics, which showed that the participants' median scores for the 2,000 VLT were arranged between 1.00 and 3.00. The median scores for both the 3,000 and 5,000 VLTs were 0.00 and 0.50, respectively. The mean scores, standard deviation, maximum and minimum scores, and range are set out in Table 4.4.

Table 4.4
Descriptive Statistics for the VLTs (28-day RI Sub-groups)

Group (N)		2,000 VLT	3,000 VLT	5,000 VLT
Control	Mean (*SD*)	3.00 (*3.02*)	1.12 (*1.65*)	0.88 (*2.15*)
(17)	Median	3.00	0.50	0.00
	Min-Max	0.00, 10.00	0.00, 5.00	0.00, 7.00
	Range	10.00	5.00	7.00
Intensive	Mean (*SD*)	2.71 (*2.87*)	1.00 (*1.77*)	0.24 (*.56*)
(17)	Median	1.00	0.00	0.00
	Min-Max	0.00, 9.00	0.00, 6.00	0.00, 2.00
	Range	9.00	6.00	2.00
Spaced	Mean (*SD*)	1.93 (*2.40*)	1.33 (*1.68*)	0.40 (*1.06*)
(15)	Median	1.00	0.00	0.00
	Min-Max	0.00, 8.00	0.00, 4.00	0.00, 3.00
	Range	8.00	4.00	3.00

Note: Maximum score for each VLT is 30.00

For the 28-day RI sub-groups, the Shapiro-Wilk test determined that the assumption of normal distribution was violated ($p < .05$). When analysing the VLT results, some outliers were found (see Figures 4.4, 4.5 & 4.6). Moreover, although homogeneity of variance was achieved in the results of the 2,000 and 3,000 VLTs, this assumption was violated in the 5,000 VLT ($p < .05$). Therefore, a Kruskal-Wallis test was carried out to compare the scores for the VLTs administered to the three groups.

Figure 4.4
Boxplot of Test Scores on the 2,000 VLT for the 28-day RI Sub-groups

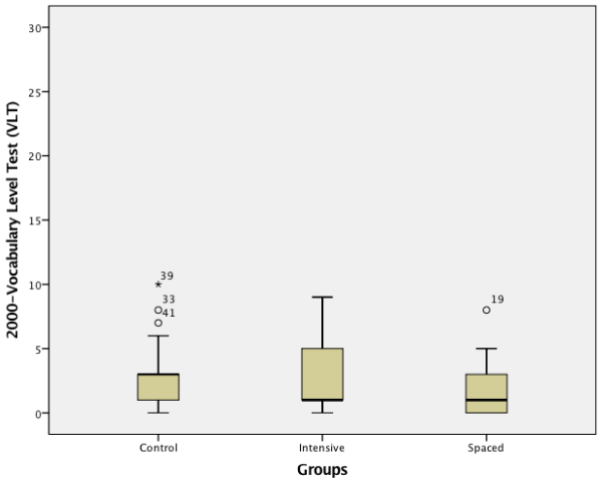


Figure 4.5
Boxplot of Test Scores on the 3,000 VLT for the 28-day RI Sub-groups

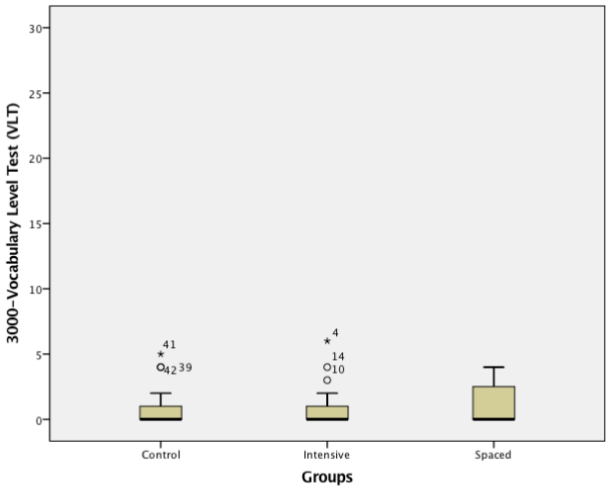
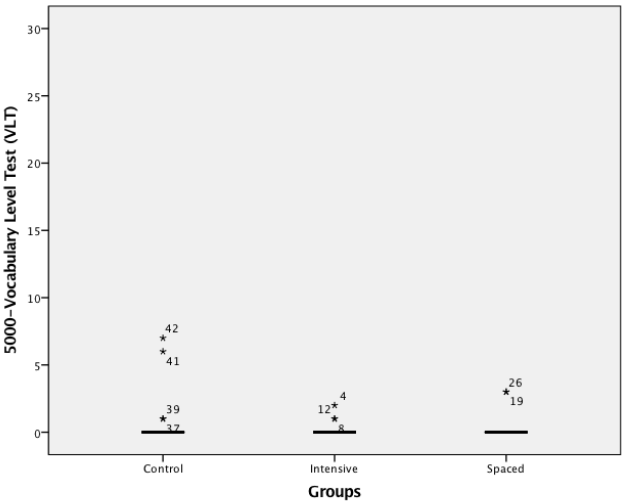


Figure 4.6
Boxplot of Test Scores on the 5,000 VLT for the 28-day RI Sub-groups



The results indicated that there was no significant difference between the three sub-groups in the results of the 2,000 VLT: $H(2)=1.68$, $p= .433$. Neither was there any statistical difference between the participants' results in the 3,000 VLT: ($H(2)= .54$, $p= .764$. Correspondingly, neither did the results of the 5,000 VLT show any significant differences between the groups: ($H(2)=.56$, $p= .755$). Therefore, none of the participants in the three main groups (within either the 4-day or 28-day RI sub-groups) exhibited any differences in existing vocabulary knowledge prior to the study and they all scored below the 2,000-word frequency level, indicating that their existing vocabulary knowledge was low.

4.4.2 Working Memory (WM) Tests

The second baseline tests applied were the WM tests (consisting of digit span tests). These included a forward and backward test, conducted for each individual student (see Appendix 11). The WM tests were administered to the participants after they had completed all of the treatment sessions, so that the test results would not interfere with the intervention. It took around two weeks to collect all the WM test data. However, a few students missed the tests, because they had dropped the college course and were no longer attending the classes. These tests were conducted after the treatment sessions to evaluate the participants' ability to recall information, which would help to establish the homogeneity of the groups.

Initially, descriptive statistics were generated to give an overview of the two WM tests for the three Groups (Control, Intensive and Spaced). As with the VLT results, the results of the WM tests for the 4-day RI and 28-day RI sub-groups were analysed separately and will be presented sequentially in the following sections.

4.4.2.1 The 4-day Retrieval Interval (RI) Sub-groups

The WM descriptive statistics for the 4-day RI sub-groups showed that the participants across all the groups achieved comparable mean and median scores in both WM tests (forward and backward), as shown in Table 4.5. Moreover, the Shapiro-Wilk test revealed that the WM data were normally distributed ($p> .05$) in the results achieved by all the 4-day RI sub-groups. A boxplot was then generated for both tests (see Figures 4.7 and 4.8), in order to find out if there were any extreme values that should be deleted, with a parametric test being run (one-

way ANOVA). In addition, the results of both WM tests, generated using Levene's test, met the assumption of homogeneity of variance in the forward WM (p= .517), and also in the backward WM (p= .485). Therefore, all the WM test scores were subjected to one-way ANOVA, as the three main assumptions were fulfilled for running the parametric test (see Table 4.5).

Table 4.5

Descriptive Statistics from the WM tests (Forward and Backward) and the Normality Test (Shapiro-Wilk) for the 4-day RI Sub-groups

Group (N)		**WM Test (Forward)**	**WM Test (Backward)**
Control	Mean (*SD*)	7.94 (*1.65*)	5.63 (*1.15*)
(16)	Median	8.00	6.00
	Min-Max	5.00, 11.00	3.00, 7.00
	Range	6.00	4.00
	P-value	.497	.069
Intensive	Mean (*SD*)	8.18 (*1.55*)	5.94 (*1.98*)
(17)	Median	8.00	6.00
	Min-Max	6.00, 11.00	3.00, 11.00
	Range	5.00	8.00
	P-value	.153	.091
Spaced	Mean (*SD*)	8.57 (*1.22*)	5.86 (*1.79*)
(14)	Median	8.50	6.00
	Min-Max	6.00, 10.00	3.00, 10.00
	Range	4.00	7.00
	P-value	.094	.091

Figure 4.7
Boxplot of WM Test Scores (Forward) (4-day RI Sub-groups)

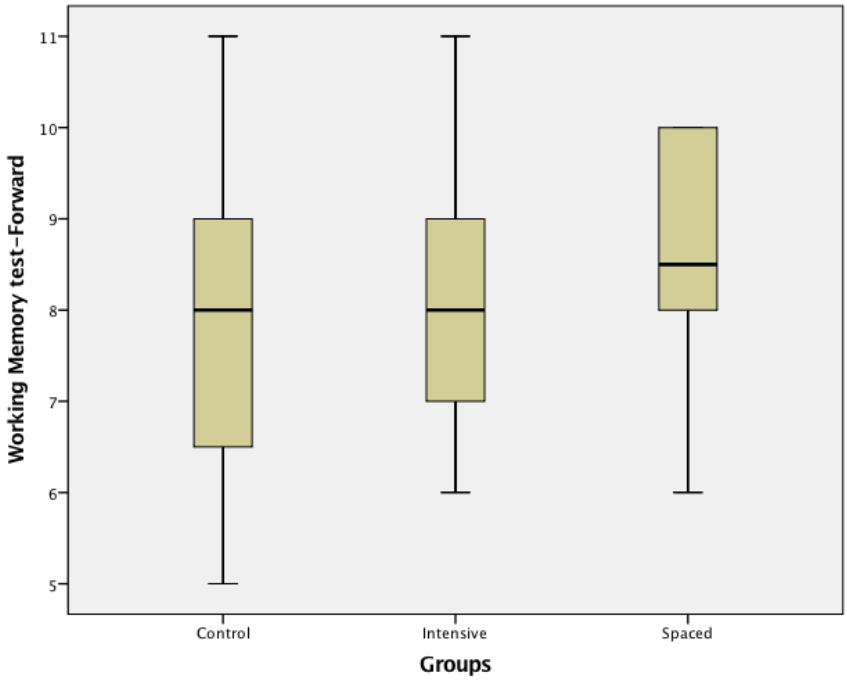
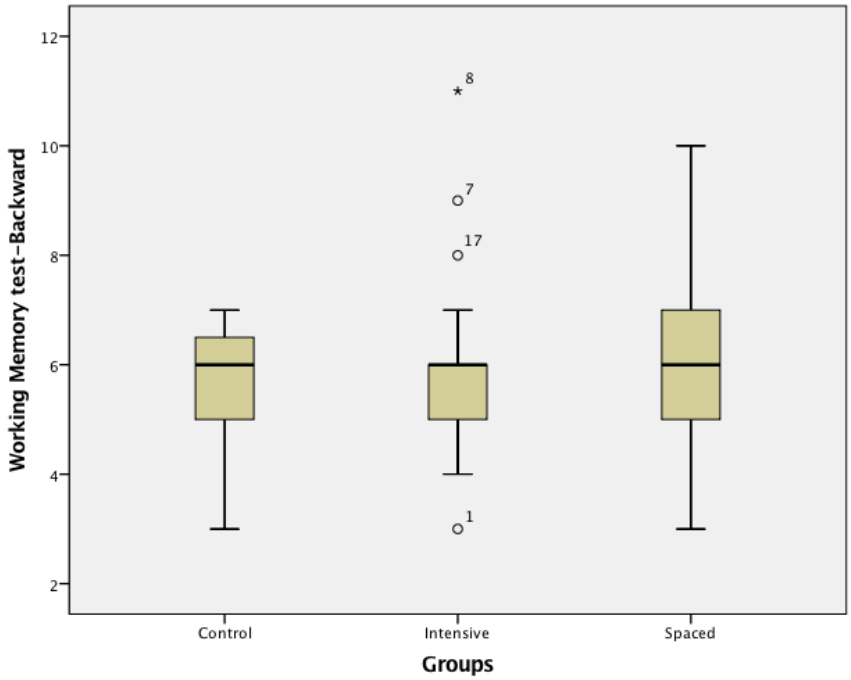


Figure 4.8
Boxplot of WM Test Scores (Backward) (4-day RI Sub-groups)



The results of the WM test (forward) generated using one-way ANOVA revealed no significant differences between the 4-day RI sub-groups (with the Control, Intensive and Spaced Groups): $F(2, 44)= .68$, $p=.514$, $\omega=0.87$. In addition, there was no significant difference in mean scores for the WM (backward) test between the same groups: $F(2, 43)= .12$, $p= .891$, $\omega=-0.14$.

To sum up, neither of the WM tests (forward or backward) revealed any statistical differences between the 4-day RI sub-groups. Therefore, it was not necessary to run a further analysis of the post-hoc test. The following section explores the data from the WM tests across the 28-day RI sub-groups.

4.4.2.2 The 28-Day Retrieval Interval (RI) sub-groups

The WM descriptive statistics for the 28-day RI sub-groups are presented in Table 4.6. The mean scores from the WM test (forward) for all groups ranked between 7.88 and 8.27, with a median of 8.00 for each group. In the WM test (backward), the results also showed comparable mean scores of 5.35, 5.82 and 6.40, and median scores of 5.00, 6.00 and 6.00, respectively, for the Control, Intensive and Spaced 28-day RI sub-groups. In addition, both WM tests (forward and backward) displayed a more or less normal distribution ($p> .05$), assessed visually by inspecting the boxplots (see Figures 4.9 and 4.10). The assumption of homogeneity of variance was satisfied ($p> .05$), checking the data from the forward WM ($p= .517$) and backward WM ($p= .485$). Therefore, both sets of WM test scores were analysed using one-way ANOVA, as they fulfilled the assumptions required to apply this parametric test (see Table 4.6).

Table 4.6

Descriptive Statistics from the WM tests (Forward and Backward) and Normality Test (Shapiro-Wilk) for the 28-day RI Sub-groups

Group (N)		WM Test (Forward)	WM Test (Backward)
Control	Mean (*SD*)	7.88 (*1.65*)	5.35 (*1.77*)
(16)	Median	8.00	5.00
	Min-Max	5.00, 10.00	3.00, 8.00
	Range	5.00	5.00
	P-value	.022	.100
Intensive	Mean (*SD*)	7.71 (*1.61*)	5.82 (*1.70*)
(17)	Median	8.00	6.00
	Min-Max	5.00, 11.00	2.00, 9.00
	Range	6.00	7.00
	P-value	.494	.677
Spaced	Mean (*SD*)	8.27 (*1.62*)	6.40 (*1.96*)
(14)	Median	8.00	6.00
	Min-Max	5.00, 11.00	2.00, 10.00
	Range	6.00	8.00
	P-value	.284	.615

Figure 4.9

Boxplot of the WM Test Scores (Forward) for the 28-day RI Sub-groups

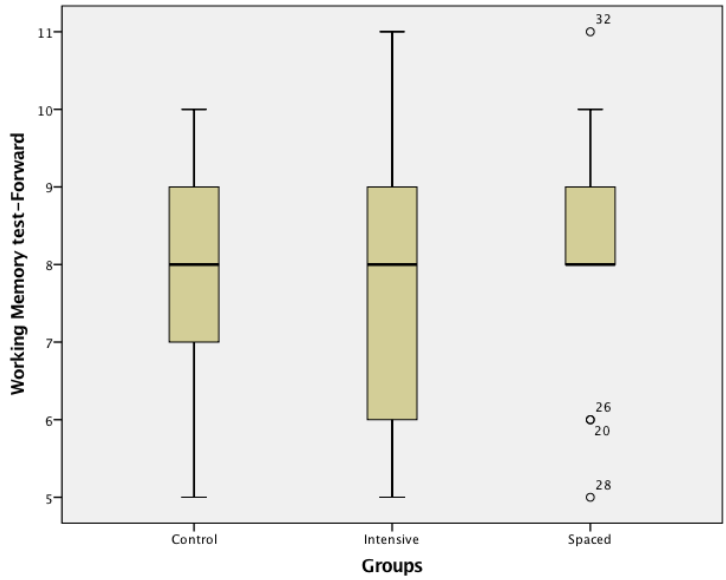
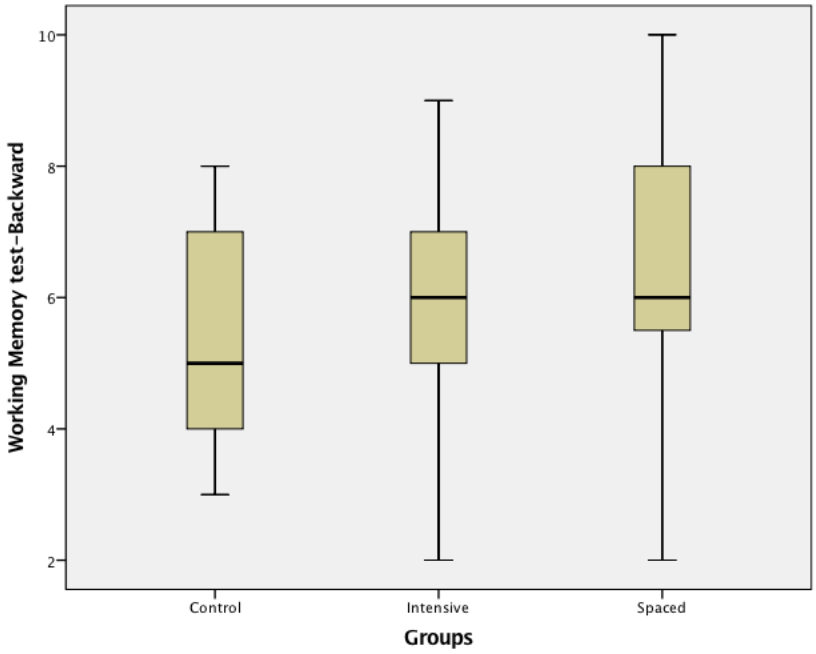


Figure 4.10

Boxplot of the WM Test Scores (Backward) for the 28-day RI Sub-groups



The results of the one-way ANOVA showed that there were no significant differences in the WM (forward) test results achieved by the Control, Intensive and Spaced 28-day RI sub-groups: $F(2, 46)= .49$, $p=.618$, $\omega= -0.02$. Similar results were obtained from the WM (backward) test, with no significant difference between the groups: $F(2, 46)=1.34$, $p= .272$, $\omega=0.18$.

Overall, there were no significant differences found between the WM test scores (forward or backward). Thus, there was no need for further analysis, as the results showed no difference in the participants' ability to retain vocabulary. Therefore, it was determined that this factor would not affect the results of the intervention, which are presented in the following section.

4.5 Experimental Vocabulary Learning Gains

The first and second research questions endeavoured to ascertain whether the three groups (Control, Intensive and Spaced) progressed differently from each other between the three time points (the pre-, immediate post-, and delayed post-tests). Therefore, a two-way mixed ANOVA was conducted to ascertain whether the three groups had progressed differently from each other over the three time-points. In addition, *post hoc* pairwise comparisons were made

using Bonferroni corrections, in order to investigate whether there had been any improvement in each group between the different test time points. As with the previous methods of analysis, the groups' results were divided into two sub-groups: 4-day RI and 28-day RI.

Initially, frequency and descriptive statistics were used to provide a full description of the participants' performance in both the Control and Experimental (Intensive and Spaced) Groups at the three different test times (pre-, immediate post-, and delayed post-tests). The outcomes of the data analysis begin here with a detailed look at the 4-day RI sub-group, and are followed by the 28-day RI sub-group in similar detail. It is worth mentioning that because some of the data were non-normally distributed, non-parametric tests were run to confirm the results of the parametric two-way ANOVA analyses (see section 3.5.3). The non-parametric analyses supported all of the results elicited from the parametric analyses (see Appendices 16, 17, 18 and 19 for the results of the non-parametric analyses). Therefore, this section will present the results of the two-way ANOVA analyses.

4.5.1 Frequency with Which the Students Attempted Answers

This section is intended to present a clear picture of the students attempting to answer questions, looking at their missing answers over the three performance tests (pre-, immediate post-, and delayed post-) for the 4-day RI and 28-day sub-groups. In addition, this section aims to measure any improvement in the students' correct answers, particularly with regard to correct spelling and meaning. Therefore, Table 4.7 and Figure 4.11 were derived from the Active Recall (translation from L1 to L2) and Passive Recall (translation from L2 to L1) sub-tests of the three main performance tests (pre-, immediate post-, and delayed post-).

Table 4.7
Percentage of Missing Answers from the 4-day RI Sub-groups

	Pre-test %	Immediate post-test %	Delayed post-test %
Control	68.2	64	63.5
Intensive	59.5	17.1	19
Spaced	62.4	20.4	19.3

Figure 4.11
Line Chart of Missing Answers for the 4-day RI Sub-groups

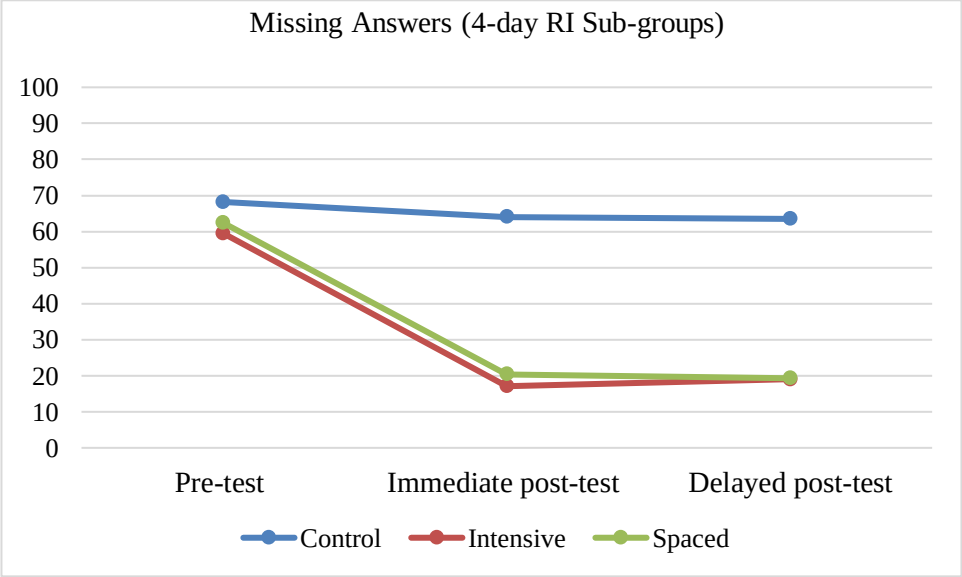


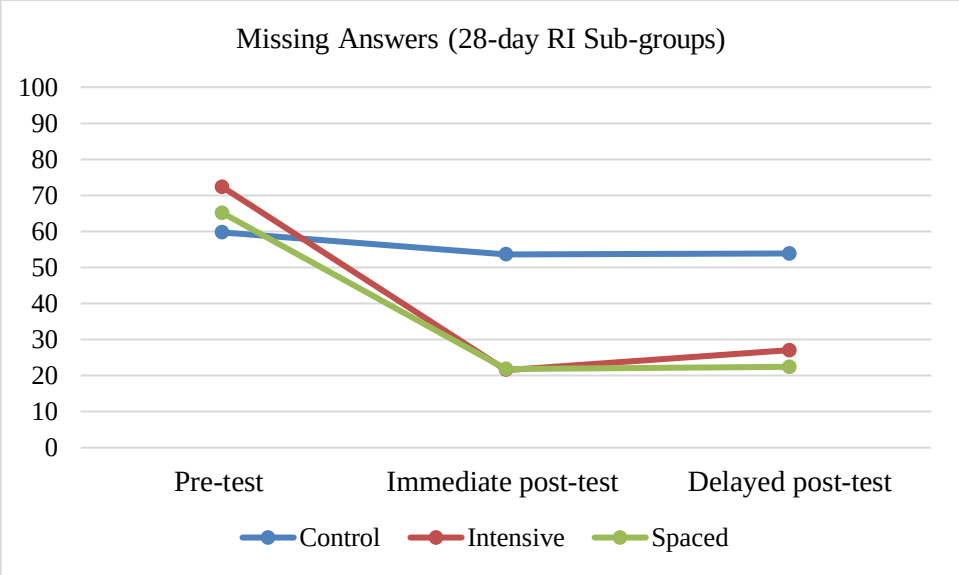
Table 4.7 and Figure 4.11 display a large decrease in the number of missing answers in the Experimental Groups (Intensive and Spaced), from the pre-test to the immediate and delayed (4-day RI) post-tests. However, the number of missing answers in the Control Group indicated no change over the three time points. This could indicate the students' motivation in the Experimental Groups to attempt more answers at post- and delayed post-test.

Table 4.8 and Figure 4.12 illustrate the students' missing answers in the 28-day RI sub-groups (Intensive, Spaced and Control).

Table 4.8
Percentage of Missing Answers from the 28-day RI Sub-groups

	Pre-test %	Immediate post-test %	Delayed post-test %
Control	59.7	53.6	53.8
Intensive	72.3	21.5	27
Spaced	65.1	21.8	22.4

Figure 4.12
Line Chart of Missing Answers for the 28-day RI Sub-groups



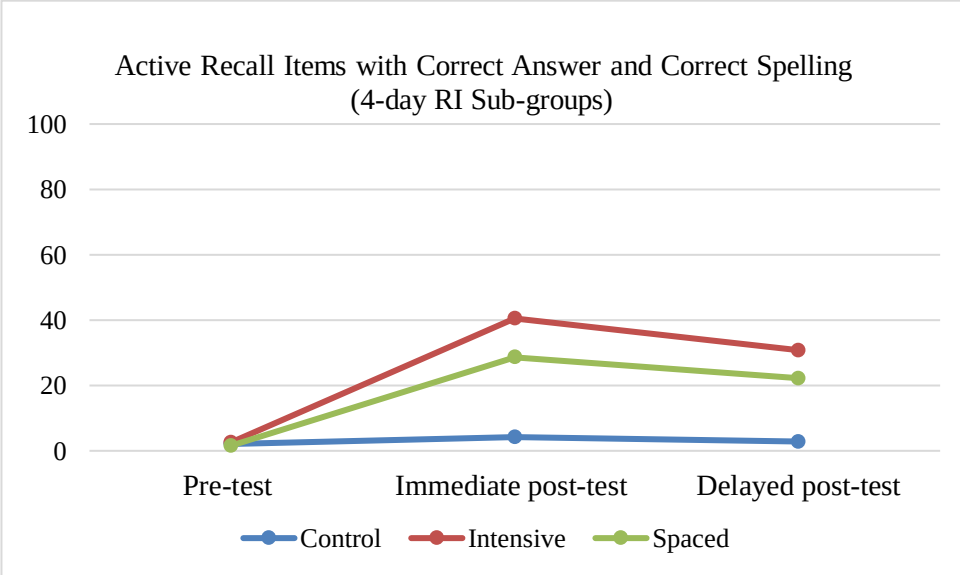
As with the 4-day RI subgroups, Table 4.8 and Figure 4.12 revealed that both Experimental Groups demonstrated a reduction in the number of missing answers in the immediate post-test, with sustainable results at the delayed (28-day RI) post-test. Conversely, the Control Group showed no change in their attempts to answer, whether in the pre-test, post-test or delayed (28-day RI) post-test. Therefore, the researcher subsequently checked the number of fully correct answers, i.e. those with the correct meaning and spelling in the same sub-tests (Passive and Active Recall) among the Experimental Groups (Intensive and Spaced) and in the Control Group. The purpose of this was to check the effectiveness of the intervention for the groups in the current study, not only for recognition of the target vocabulary, but also ability to accurately recall the vocabulary items. This means that the score results in this section are based on the students' identification of the correct word

without allowing any spelling errors. However, when running the inferential statistics in the following sections, the researcher followed the College way of correction, which is allowing fewer spelling errors (see section 3.5.1).

Table 4.9
Percentage of Active Recall Correct Answers from the 4-day RI Sub-groups

	Pre-test %	Immediate post-test %	Delayed post-test %
Control	2.1	4.2	2.8
Intensive	2.6	40.5	30.7
Spaced	1.6	28.6	22.2

Figure 4.13
Line Chart of Active Recall Correct Answers from the 4-day RI Sub-groups



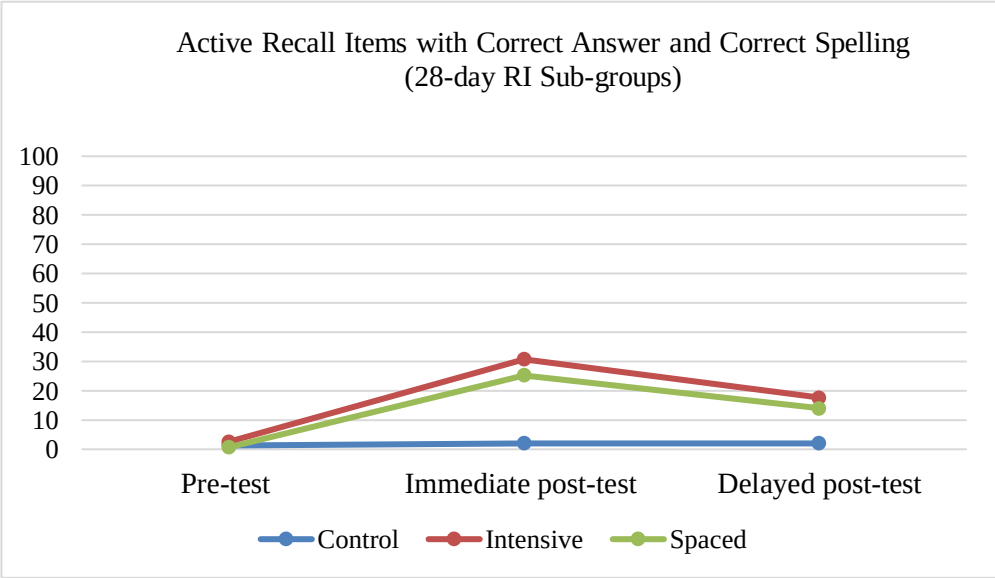
The results of Active Recall (translation from L1 into L2) from the 4-day RI sub-groups indicate a major increase in the number of correct answers from the Intensive and Spaced Groups, from the pre-test to the immediate post- and delayed post-tests. Meanwhile, the Control Group did not record any improvement from the pre-test to the delayed post-test (see Table 4.9 and Figure 4.13). This result indicates the positive impact of the intervention on the students in the Experimental Groups because of the students' accuracy in producing entirely correct spelling at the immediate test, with very slight decline after the short-term retrieval

interval (4-day RI). However, the results from this Active Recall task also showed some differences between both Experimental Groups with slightly better performance by the Intensive Group at immediate and delayed (4-day RI) post-tests.

Table 4.10
Percentage of Active Recall Correct Answers from the 28-day RI Sub-groups

	Pre-test %	Immediate post-test %	Delayed post-test %
Control	1.3	2	2
Intensive	2.6	30.7	17.6
Spaced	0.7	25.2	14

Figure 4.14
Line Chart of Active Recall Correct Answers from the 28-day RI Sub-groups



Similarly, there was a considerable improvement in producing the form and retention of the meaning of the target words among the Experimental 28-day RI sub-groups, who were found to have made great progress in the immediate post-test. They were further found to have maintained this progress with a slight drop by the delayed (28-day RI) post-test. However, the Control Group demonstrated no improvement or change in their results across the three performance sub-tests (see Table 4.10 and Figure 4.14). This reflects that both Experimental Groups were able to do well in the Active Recall through producing the word

(form and meaning) after the long-term retrieval interval, with no differences apparent between both sub-groups (Intensive and Spaced).

Turning to Passive Recall knowledge (translation from L2 into L1), Table 4.11 and Figure 4.15 present the students' performance.

Table 4.11
Percentage of Passive Recall Correct Answers from the 4-day RI Sub-groups

	Pre-test %	Immediate post-test %	Delayed post-test %
Control	3.1	4	0.8
Intensive	4.4	69.1	62.5
Spaced	2.7	67	67

Figure 4.15
Line Chart of Passive Recall Correct Answers from the 4-day RI Sub-groups

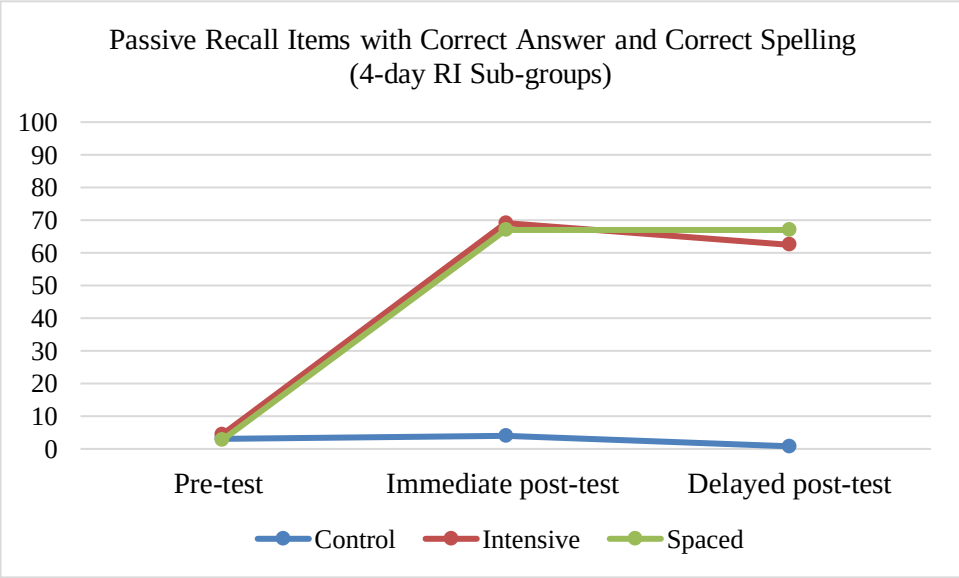
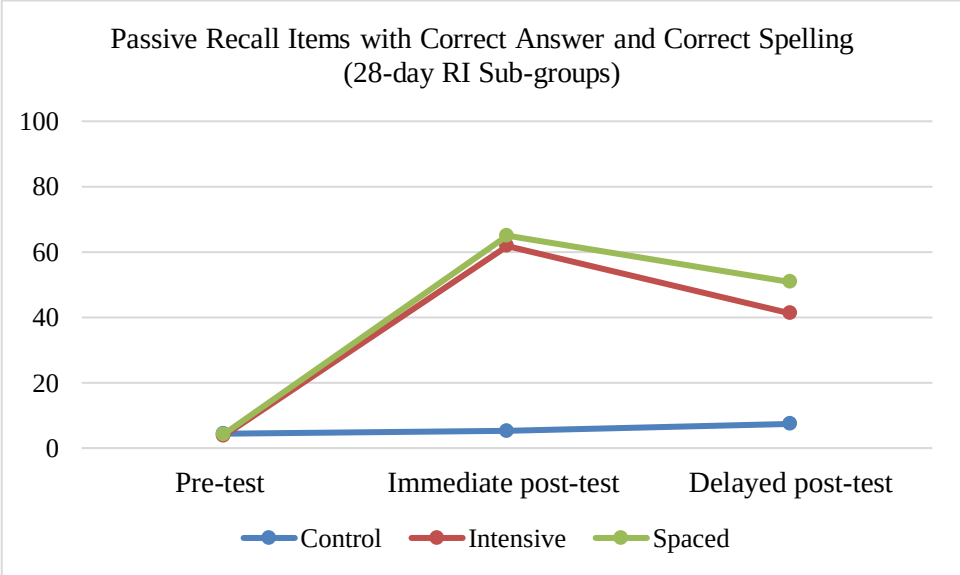


Table 4.11 and Figure 4.15 indicate that in respect of the Passive Recall results, the Experimental 4-day RI sub-groups showed remarkable progress, compared with the Control Group at the immediate post- and delayed post-tests. This progress thereby reflects the benefits of the practice sessions in short-term retrieval for recognising the word meanings and translating them into Arabic equivalents.

Table 4.12
Percentage of Passive Recall Correct Answers from the 28-day RI Sub-groups

	Pre-test %	Immediate post-test %	Delayed post-test %
Control	4.4	5.2	7.4
Intensive	3.7	61.8	41.2
Spaced	4.2	65	50.8

Figure 4.16
Line Chart of Passive Recall Correct Answers from the 28-day RI Sub-groups



Regarding the 28-day RI sub-groups, they showed significant progress from the Experimental Groups, with no progress for the Control Group in both post-tests. In the delayed (28-day RI) post-test, the score results of the Experimental Group dropped slightly. However, the performance of the Spaced Group was better than the Intensive (see Table 4.12 and Figure 4.16).

To sum up, both Experimental Groups showed great progress at the immediate post-test, which they maintained in the shorter and longer delayed post-tests. The results of Active Recall (translation from L1 into L2) from the delayed (4-day RI) post-test showed a bias toward the Intensive Group, while the results of the Passive Recall (translation from L2 into L1) from the delayed (28-day RI) post-test indicated the slight outperformance of the Spaced

Group. Therefore, the intervention practice sessions received by the Experimental Groups were extremely beneficial for improving Passive Recall knowledge, and it was also useful to enhance the Active Recall knowledge for short- and long-term retention.

4.5.2 Pre-, Immediate Post-, and Delayed Post-tests (4-day RI Sub-groups)

This section focuses on the results of the two main categories (Recognition and Recall), as seen in the descriptive statistical data in Table 4.13.

Table 4.13
Descriptive Statistical Data for Each Test Time Point (4-day RI Sub-groups)

Group (N)		**Recognition**			**Recall**		
		Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)	Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)
Control	Mean (*SD*)	5.31 (*3.79*)	5.50 (*4.05*)	6.75 (*3.82*)	0.56 (*1.55*)	0.88 (*1.71*)	0.63 (*1.31*)
(16)	95% CI	[3.29, 7.33]	[3.34, 7.66]	[4.71, 8.79]	[-0.26, 1.39]	[-0.04, 1.79]	[-0.07, 1.32]
	Min-Max	0.00-13.00	0.00-14.00	0.00-15.00	0.00-6.00	0.00-6.00	0.00-5.00
Intensive	Mean (*SD*)	5.65 (*2.78*)	15.53 (*2.94*)	15.12 (*3.69*)	1.18 (*1.47*)	10.59 (*4.61*)	9.47 (*4.60*)
(17)	95% CI	[4.22, 7.08]	[14.02, 17.04]	[13.22, 17.01]	[0.42, 1.93]	[8.22, 12.96]	[7.11, 11.83]
	Min-Max	1.00-12.00	8.00-17.00	5.00-17.00	0.00-4.00	3.00-17.00	1.00-17.00
Spaced	Mean (*SD*)	5.57 (*3.37*)	16.07(*1.49*)	15.71 (*1.77*)	0.64 (*1.08*)	9.50 (*4.75*)	9.43 (*4.52*)
(14)	95% CI	[3.63, 7.52]	[15.21, 16.93]	[14.69, 16.74]	[0.02, 1.27]	[6.76, 12.24]	[6.82, 12.04]
	Min-Max	0.00-14.00	12.00-17.00	11.00-17.00	0.00-4.00	1.00-17.00	1.00-17.00

Note: Maximum test scores of 18.00 for Recall tasks and 16.00 for Recognition tasks

The descriptive statistics in Table 4.13 signify that the Experimental Groups (Intensive and Spaced) made progress in the immediate post-test and maintained this progress in the delayed post-test, evidenced by the means for both categories of question (Recognition and Recall). In contrast, there was no change in the Control Group's mean scores for Recognition and Recall, achieved over the three time points (pre-, immediate post-, and delayed post-tests).

4.5.2.1 Recognition

The scores for the Recognition items across the three tests (pre-, immediate post-, and delayed post-) for the 4-day RI sub-groups were subsequently submitted to a two-way mixed ANOVA. The results indicated a significant main effect of Time: $F(1.6, 88)=206.9$, $p< .001$, $\eta_{p2}= .825$, $\eta_2= .618$. The main Group effect was also statistically significant: $F(2, 44)=27.04$, $p< .001$, $\eta_{p2}= .551$, $\eta_2= .349$. Furthermore, the Group * Time interaction was statistically significant: $F(3.18, 88)=41.84$, $p< .001$, $\eta_{p2}= .655$, $\eta_2= .250$. This interaction reflects statistically significant differences in the learning trajectories of the three groups over time, which is reflected in the large effect sizes (Table 4.14) between Time 1 and Time 2 / Time 3 for the Experimental Groups but very small effect sizes for the Control Group.

Table 4.14
Effect Sizes for Repeated Measures (within-subjects) Comparison between the Different Test Time Points for Each Group (4-day RI Sub-groups)

Group		Time 2	Time 3
Control	**Time 1**	*p*=1.00, *d*=0.05, 95% CI [-0.93, 1.03]	*p*= .314, *d*=0.38, 95% CI [-0.61, 1.37]
	Time 2		*p*< .035, *d*=0.32, 95% CI [-0.67, 1.30]
Intensive	**Time 1**	*p*< .001, *d*=3.45, 95% CI [1.95, 4.95]	*p*< .001, *d*=2.90, 95% CI [1.54, 4.26]
	Time 2		*p*=1.00, *d*= -0.12, 95% CI [-1.08, 0.83]
Spaced	**Time 1**	*p*< .001, *d*=4.03, 95% CI [2.21, 5.85]	*p*< .001, *d*=3.77, 95% CI [2.02, 5.51]
	Time 2		*p*=1.00, *d*= -0.22, 95% CI [-1.27, 0.83]

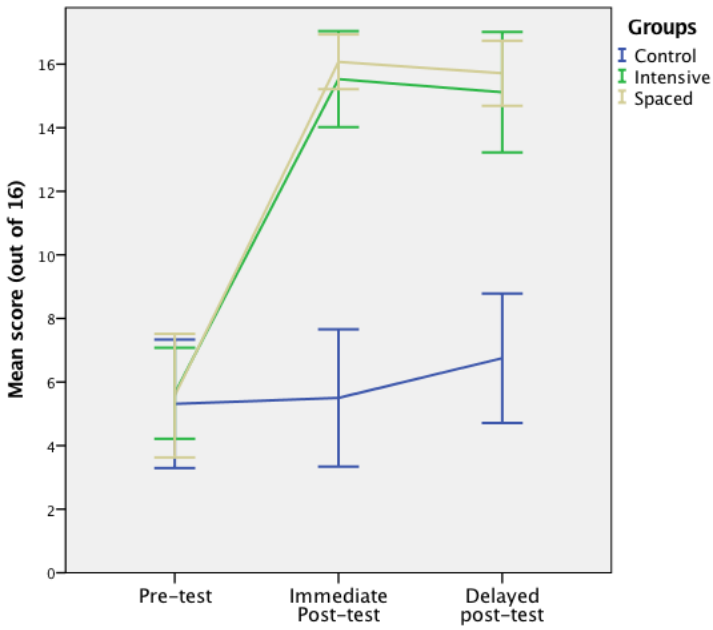
Post hoc pairwise comparisons (see Table 4.14) were also carried out using Bonferroni corrections, which generated results to show that significant progress was made by both the Intensive and Spaced Groups between Time 1 and Time 2, and between Time 1 and Time 3. This improvement is reflected in the very large effect sizes (Table 4.14), with confidence intervals that do not cross zero, suggesting that these are reliable effects. However, the effect size and confidence intervals indicate that the Control Group did not make any significant progress between Time 1 and Time 2, *d*= .05 (-.93, 1.03). Neither were there any significant differences displayed by the Experimental Groups (Intensive and Spaced) between Time 2

and 3, suggesting that both groups maintained their higher scores at delayed post-test. In addition, there was a significant change from Time 2 to Time 3 for the Control Group, but as noted in Table 4.14, the effect size indicates this was very small, with a 95% confidence interval crossing zero (Plonsky & Oswald, 2014). Nevertheless, it should be noted that this could reflect a small test effect, as there were only four days between Times 2 and 3. Table 4.15 presents the effect sizes for the between group comparisons in detail.

Table 4.15
Between-subjects Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (4-day Sub-groups)

Group		Intensive	Spaced
Time 1	**Control**	*p*= .946, *d*=0.10, 95% CI [-0.58, 0.79]	*p*=1.00, *d*=0.07, 95% CI [-0.65, 0.79]
	Intensive		*p*=1.00, *d*=-0.03, 95% CI [-0.73, 0.68]
Time 2	**Control**	*p*< .001, *d*=2.85, 95% CI [1.88, 3.82]	*p*< .001, *d*=3.37, 95% CI [2.26, 4.49]
	Intensive		*p*=1.00, *d*=0.23, 95% CI [-0.49, 0.94]
Time 3	**Control**	*p*< .001, *d*=2.23, 95% CI [1.36, 3.10]	*p*< .001, *d*=2.94, 95% CI [1.91, 3.98]
	Intensive		*p*=1.00, *d*=0.20, 95% CI [-0.51, 0.91]

Figure 4.17
Mean Scores for the Vocabulary Recognition Task over Time and by Group (4-day RI Sub-groups)



The results of comparisons between groups, illustrated Table 4.15 and Figure 4.17, also expose that the Control Group's performance was significantly lower than the two Experimental Groups at both immediate post-test and delayed post-test, and this is reflected in the large effect sizes (see Table 4.15). The Intensive and Spaced Groups produced equally satisfactory results, with no differences at the immediate and delayed (4-day RI) post-tests.

4.5.2.2 Recall

The Recall results indicated that there was a significant main effect of Time: $F(1.16, 88)=144.66$, $p< .001$, $\eta_{p2}= .767$, $\eta_2= .562$, as well as Group: $F(2, 44)=25.04$, $p< .001$, $\eta_{p2}= .532$, $\eta_2= .233$. Additionally, the Group * Time interaction was statistically significant: $F(2.31, 88)=34.68$, $p< .001$, $\eta_{p2}= .612$, $\eta_2= .269$, reflecting the fact that there were significant differences between the Control and Experimental Groups over the three time points, supported by the large effect sizes. Table 4.16 displays no significant differences for the Control Group between Time 1 and Time 2, or between Time 1 and Time 3. Conversely, both the Intensive and Spaced Groups made significant progress between Time 1 and Time 2, and between Time 1 and Time 3, and this is reflected in the very large effect sizes (Table 4.16). Additionally, the confidence intervals do not cross zero, which suggest that these results are reliable effects. In contrast, between Times 2 and 3, no significant differences were noted for the Spaced Group, whereas a significant drop was recorded for the Intensive Group, albeit with a very small effect, and a confidence interval crossing zero, which suggests that this effect was unreliable.

Table 4.16
Repeated Measures (within-subjects) Comparison between the Different Test Time Points for Each Group on the Recall Test (4-day RI Sub-groups)

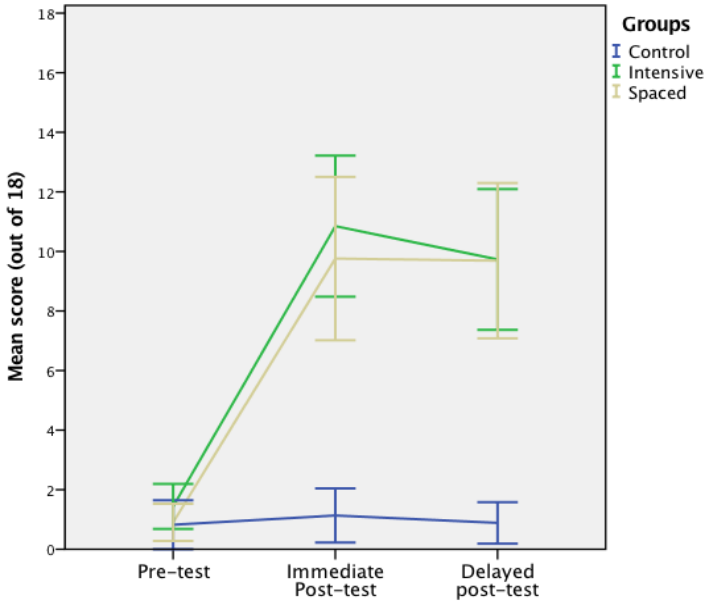
Group		Time 2	Time 3
Control	**Time 1**	*p*=1.00, *d*=0.20, 95% CI [-0.79, 1.18]	*p*=1.00, *d*=0.05, 95% CI [-0.93, 1.03]
	Time 2		*p*=1.00, *d*= -0.16, 95% CI [-1.15, 0.82]
Intensive	**Time 1**	*p*< .001, *d*=2.75, 95% CI [1.42, 4.08]	*p*< .001, *d*=2.43, 95% CI [1.18, 3.68]
	Time 2		*p*< .001, *d*= -0.24, 95% CI [-1.20, 0.71]
Spaced	**Time 1**	*p*< .001, *d*=2.57, 95% CI [1.16, 3.99]	*p*< .001, *d*=2.68, 95% CI [1.23, 4.12]
	Time 2		*p*=1.00, *d*= -0.02, 95% CI [-1.06, 1.03]

Table 4.17
Between-subjects Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point on the Recall test (4-day Sub-groups)

Group		Intensive	Spaced
Time 1	**Control**	*p*=1.00, *d*=0.05, 95% CI [-0.93, 1.03]	*p*=1.00, *d*=0.06, 95% CI [-0.66, 0.78]
	Intensive		*p*=1.00, *d*= -0.41, 95% CI [-1.13, 0.30]
Time 2	**Control**	*p*< .001, *d*=2.76, 95% CI [1.81, 3.71]	*p*< .001, *d*=2.48, 95% CI [1.53, 3.44]
	Intensive		*p*=1.00, *d*= -0.23, 95% CI [-0.94, 0.48]
Time 3	**Control**	*p*< .002, *d*=2.58, 95% CI [1.66, 3.50]	*p*< .001, *d*=2.73, 95% CI [1.73, 3.72]
	Intensive		*p*=1.00, *d*= -0.01, 95% CI [-0.72, 0.70]

Figure 4.18

Mean Scores for the Vocabulary Recall Task over Time and by Group (4-day RI Sub-groups)



The results presented in Table 4.17 and Figure 4.18 illustrate that both the Intensive and Spaced Groups improved significantly, with no differences between the groups at immediate post-test. In addition, both groups had virtually the same score at delayed (4-day RI) post-test. However, at delayed (4-day RI) post-test, the Spaced Group's retention was better (i.e. there was no decrease in their scores), as illustrated in Figure 4.18. The performance of the Intensive Group appeared to have diminished somewhat at delayed post-test, although the effect was very small and unreliable, *d*= -0.24, (-1.20, 0.71) (see Table 4.17). In contrast, the Control Group displayed no significant change over time, either in the short (at immediate post-test) or long term (at delayed post-test). Furthermore, the Control Group's scores were significantly lower than both the Intensive and Spaced Groups' at immediate post-test and delayed (4-day RI) post-test, noting that the effect sizes for the between group comparisons were very large and reliable (see Table 4.17).

4.5.3 Pre-, Immediate Post-, and Delayed Post-tests (28-day RI Sub-groups)

First, the descriptive data statistics (see Table 4.18) for the 28-day RI sub-groups, with respect to the Control, Intensive and Spaced Groups, are presented here as an overview of the students' performance. As shown in Table 4.18, the descriptive data represent the results of the students' performance in the two main test categories (Recall and Recognition) at each

time point. Section 3.4.4.4 in Chapter Three gives more detail on the distribution of items in each category and section.

Table 4.18
Descriptive Data Statistics for Each Test Time Point for Each Treatment Condition (28-day RI Sub-groups)

Group (N)		Recognition			Recall		
		Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)	Pre-test	Immediate Post-test	Delayed Post-test (28-day RI)
Control (17)	Mean (*SD*)	6.71 (*3.75*)	6.88 (*4.11*)	6.88 (*4.20*)	0.76 (*1.60*)	1.53 (*2.43*)	1.06 (*1.98*)
	95% CI	[4.78, 8.64]	[4.77, 8.99]	[4.72, 9.04]	[-0.06, 1.59]	[0.28, 2.78]	[0.04, 2.08]
	Min-Max	0.00-13.00	0.00-14.00	0.00-15.00	0.00-6.00	0.00-9.00	0.00-7.00
Intensive (17)	Mean (*SD*)	5.35 (*3.61*)	16.06 (*1.35*)	15.12 (*2.55*)	0.59 (*1.00*)	8.94 (*6.15*)	6.41 (*5.19*)
	95% CI	[3.50, 7.21]	[15.37, 16.75]	[13.81, 16.43]	[0.07, 1.10]	[5.78, 12.10]	[3.75, 9.08]
	Min-Max	0.00-13.00	12.00-17.00	10.00-17.00	0.00-3.00	0.00-17.00	0.00-16.00
Spaced (15)	Mean (*SD*)	6.07 (*4.32*)	15.27 (*3.79*)	14.93 (*3.20*)	0.87 (*1.30*)	9.13 (*5.04*)	7.27 (*5.16*)
	95% CI	[3.68, 8.46]	[13.17, 17.36]	[13.16, 16.70]	[0.15, 1.59]	[6.34, 11.92]	[4.41, 10.12]
	Min-Max	0.00-14.00	2.00-17.00	8.00-17.00	0.00-4.00	2.00-17.00	0.00-16.00

Note: Maximum test scores are 18.00 for Recall tasks and 16.00 for Recognition tasks

The mean scores for Recognition, reported in the descriptive data, showed that the 28-day RI sub-groups within the Intensive and Spaced Groups made progress at immediate post-test, which was sustained in the delayed (28-day RI) post-test. Moreover, in the results for Recall, both the Intensive and Spaced Groups seemed to have improved by the immediate post-test, but there was a slight decrease in their results for Recall in the delayed post-test. The 28-day RI sub-group within the Control Group did not display any differences in either the Recognition or Recall test categories over the three time points (pre-, immediate post-, and delayed post-test). The performance of the three groups was checked using a two-way mixed ANOVA with *post hoc* pairwise comparisons with Bonferroni corrections, in order to detect any significant differences between the groups and any significant changes over time.

4.5.3.1 Recognition

The results for Recognition indicated a main effect of Time: $F(1.99, 92)=145.88$, $p< .001$, $\eta_{p2}= .760$, $\eta_2= .552$. The main Group effect was also statistically significant: $F(2, 46)=16.76$, $p< .001$, $\eta_{p2}= .422$, $\eta_2= .275$. Finally, the Group * Time interaction was likewise found to be statistically significant, $F(3.97, 92)=36.09$, $p< .001$, $\eta_{p2}= .611$ $\eta_2= .273$. Additionally, the results drawn from *post hoc* pairwise comparisons and the corresponding effect sizes (Table 4.19) indicated that both the Intensive and Spaced Groups had significantly progressed between Time 1 and Time 2, and between Time 1 and Time 3. However, the Control Group did not appear to have made any significant progress between these test times. Moreover, there were no significant differences detected between Times 2 and 3 among the Experimental Groups (Intensive and Spaced), suggesting that both groups sustained their learning gains at delayed post-test (28-day RI), despite the fact that there was a small decrease in their scores. These results are reported in detail in Table 4.19.

Table 4.19

Repeated Measures (within-subjects) Comparison between the Different Test Time Points for Each Group (28-day RI Sub-groups)

Group		Time 2	Time 3
Control	**Time 1**	*p*=1.00, *d*=0.04, 95% CI [-0.91, 0.99]	*p*=1.00, *d*=0.04, 95% CI [-0.91, 0.99]
	Time 2		*p*=1.00, *d*=0, 95% CI [-0.95, 0.95]
Intensive	**Time 1**	*p*<.001, *d*=3.93, 95% CI [2.30, 5.56]	*p*< .001, *d*=3.13, 95% CI [1.71, 4.54]
	Time 2		*p*= .604, *d*= -0.46, 95% CI [-1.42, 0.50]
Spaced	**Time 1**	*p*< .001, *d*=2.26, 95% CI [0.97, 3.56]	*p*< .001, *d*=2.33, 95% CI [1.02, 3.64]
	Time 2		*p*=1.00, *d*= -0.10, 95% CI [-1.11, 0.92]

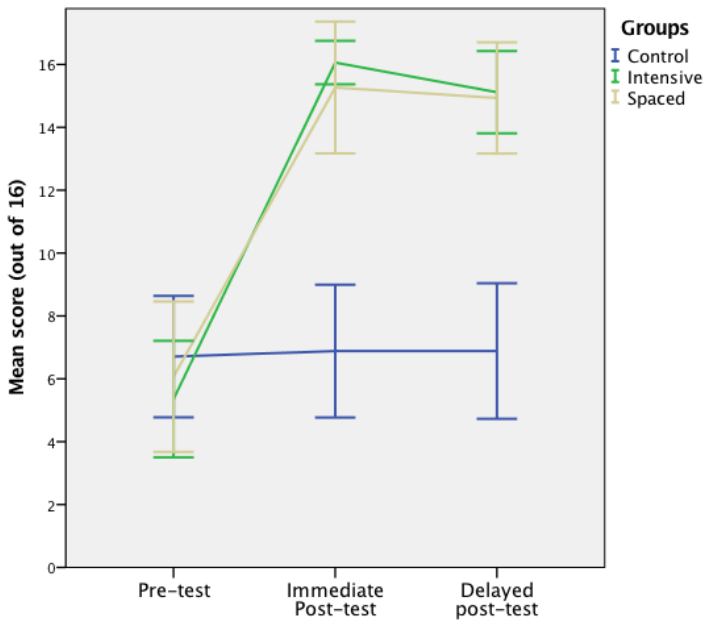
Table 4.20

Between-subjects Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (28-day RI Sub-groups)

Group		Intensive	Spaced
Time 1	**Control**	*p*= .946, *d*= -0.37, 95% CI [-1.05, 0.31]	*p*=1.00, *d*= -0.16, 95% CI [-0.85, 0.54]
	Intensive		*p*=1.00, *d*= -0.18, 95% CI [-0.51, 0.88]
Time 2	**Control**	*p*< .001, *d*=3.00, 95% CI [2.02, 3.98]	*p*< .001, *d*=2.12, 95% CI [1.25, 2.98]
	Intensive		*p*=1.00, *d*= -0.29, 95% CI [-0.98, 0.41]
Time 3	**Control**	*p*< .001, *d*=2.37, 95% CI [1.49, 3.25]	*p*< .001, *d*=2.14, 95% CI [1.27, 3.01]
	Intensive		*p*=1.00, *d*= -0.07, 95% CI [-0.76, 0.63]

Figure 4.19

Mean Scores for Vocabulary Recognition Task over Time and by Group (28-day RI Sub-groups)



Both the Intensive and Spaced Groups revealed a significant development in the students' mean scores, with no differences at any time points. However, in the delayed (28-day RI) post-test there appeared to have been a very slight drop in the results obtained by the Intensive Group (see Figure 4.19), although the effect size remained small and unreliable (the confidence interval crossed zero). In contrast, the Control Group performed significantly lower than the other groups at both immediate and delayed post-test, and this was revealed in the large between-group effect sizes (see Table 4.20).

4.5.3.2 Recall

The Recall data revealed a significant main effect of Time: $F(1.45, 92)=70.12$, $p< .001$, $\eta_{p2}= .604$, $\eta_2= .486$, and the main Group effect was also statistically significant: $F(2, 46)=10.47$, $p< .001$, $\eta_{p2}= .313$, $\eta_2= .266$. Likewise, the Group * Time interaction was found to be statistically significant: $F(2.90, 92)=14.01$, $p< .001$, $\eta_{p2}= .378$, $\eta_2= .194$. Comparisons between the three test time points were then investigated for each group, with results that showed both the Intensive and Spaced Groups progressed significantly between Time 1 and Time 2, and between Time 1 and Time 3 (Table 4.21). Meanwhile, the Control Group did not appear to have made any significant improvement between these different time points. Although there was a significant decline in performance in both the Intensive and Spaced Groups between Times 2 and 3, the effect sizes were small, and the confidence intervals crossed zero, suggesting that these effects were unreliable. These results are presented in detail in Table 4.21.

Table 4.21

Repeated Measures (within-subjects) Comparison between the Different Test Time Points for Each Group (28-day RI Sub-groups)

Group		Time 2	Time 3
Control	**Time 1**	p=1.00, d=0.37, 95% CI [-0.59, 1.33]	p=1.00, d=0.17, 95% CI [-0.79, 1.12]
	Time 2		p=1.00, d= -0.21, 95% CI [-1.17, 0.74]
Intensive	**Time 1**	p< .001, d=1.90, 95% CI [0.75, 3.04]	p< .001, d=1.56, 95% CI [0.47, 2.64]
	Time 2		p< .001, d= -0.45, 95% CI [-1.41, 0.52]
Spaced	**Time 1**	p< .001, d= 2.24, 95% CI [0.95, 3.54]	p< .001, d=1.70, 95% CI [0.52, 2.88]
	Time 2		p< .010, d= -0.37, 95% CI [-1.39, 0.66]

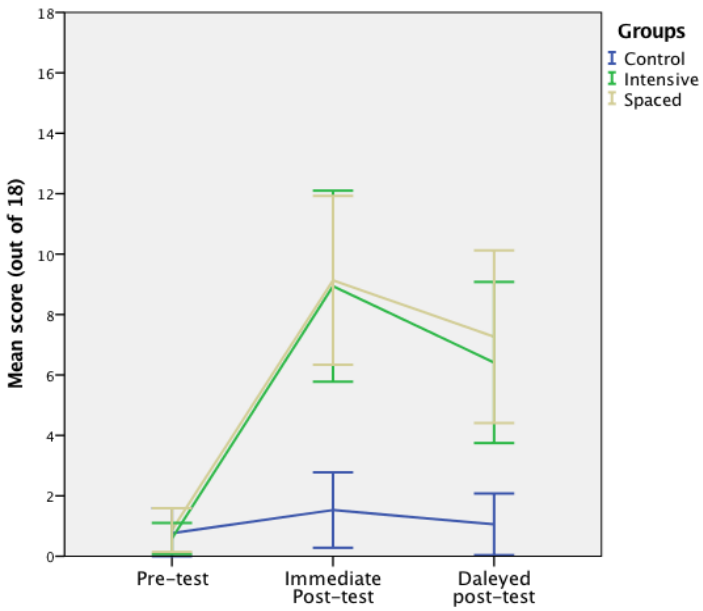
Table 4.22

Between-subjects Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (28-day RI Sub-groups)

Group		Intensive	Spaced
Time 1	**Control**	p=1.00, d= -0.13, 95% CI [-0.8, 0.55]	p=1.00, d=0.08, 95% CI [-0.62, 0.77]
	Intensive		p=1.00, d= -0.24, 95% CI [-0.45, 0.94]
Time 2	**Control**	p< .001, d=1.59, 95% CI [0.81, 2.36]	p< .001, d=1.96, 95% CI [1.12, 2.81]
	Intensive		p=1.00, d=0.03, 95% CI [-0.66, 0.73]
Time 3	**Control**	p< .002, d=1.36, 95% CI [0.62, 2.11]	p< .001, d=1.63, 95% CI [0.83, 2.43]
	Intensive		p=1.00, d= -0.17, 95% CI [-0.53, 0.86]

Figure 4.20

Mean Scores for the Vocabulary Recall Task over Time and by Group (28-day RI Sub-groups)



The between-subjects analysis presented in Table 4.22 revealed there were no significant differences between the Experimental Groups (Intensive and Spaced) at any time points (pre-, immediate and delayed tests at 28-day RI). At immediate and delayed (28-day RI) post-test, both Experimental Groups outperformed the Control Group, with significant differences and large effect sizes. The Control Group did not show any significant progress between the three different times, as clearly shown in Figure 4.20.

The following sections look at the impact of the distribution practice on the students' passive and active knowledge at the three time points (pre-, immediate and delayed) of the two main sub-groups (4-day RI and 28-day RI). A descriptive analysis will be presented for each test, focusing on passive and active knowledge. Then, the comparison between groups will take place to find out if there were any differences between the groups (Control, Intensive and Spaced).

4.5.4 The 4-day RI Sub-groups: Passive and Active Vocabulary Knowledge

This section focuses on the results of the two sub-tests (Passive and Active) in each main category, Recognition and Recall, for the 4-day RI sub-groups (Control, Intensive and Spaced), as presented in the descriptive statistical data in Table 4.23 (Passive/Active Recognition Knowledge) and Table 4.24 (Passive/Active Recall Knowledge).

Table 4.23

Descriptive Statistical Data for Recognition Knowledge (Passive/ Active) at Each Test Time Point (4-day RI Sub-groups)

Group (N)		**Recognition Knowledge**					
		Passive			**Active**		
		Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)	Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)
Control (16)	Mean (*SD*)	2.00 (*1.86*)	2.00 (*1.67*)	2.56 (*1.55*)	3.31 (*2.39*)	3.50 (*2.68*)	4.19 (2.54)
	95% CI	[1.01, 2.99]	[1.11, 2.89]	[1.74, 3.39]	[2.04, 4.58]	[2.07, 4.93]	[2.84, 5.54]
	Min-Max	0.00-6.00	0.00-6.00	0.00-6.00	0.00-9.00	0.00-8.00	0.00-9.00
Intensive (17)	Mean (*SD*)	2.00 (*1.28*)	7.18 (*1.74*)	7.12 (*2.15*)	3.65 (*2.26*)	8.35 (*1.27*)	8.00 (*1.87*)
	95% CI	[1.34, 2.66]	[6.28, 8.07]	[6.01, 8.22]	[2.48, 4.81]	[7.70, 9.01]	[7.04, 8.96]
	Min-Max	0.00-5.00	3.00-8.00	1.00-8.00	1.00-9.00	5.00-9.00	4.00-9.00
Spaced (14)	Mean (*SD*)	1.93 (*1.54*)	7.64 (*.84*)	7.43 (*1.02*)	3.64 (*2.31*)	8.43 (*.85*)	8.29 (*.99*)
	95% CI	[1.04, 2.82]	[7.16, 8.13]	[6.84, 8.02]	[2.31, 4.98]	[7.94, 8.92]	[7.71, 8.86]
	Min-Max	0.00-5.00	5.00-8.00	5.00-8.00	0.00-9.00	7.00-9.00	6.00-9.00

Table 4.24
Descriptive Statistical Data for Recall Knowledge (Passive/ Active) at Each Test Time Point (4-day RI Sub-groups)

Group (N)		Recall Knowledge					
		Passive			Active		
		Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)	Pre-test	Immediate Post-test	Delayed Post-test (4-day RI)
Control (16)	Mean (*SD*)	0.25 (*.78*)	0.31 (*.87*)	0.13 (*.50*)	0.25 (*.58*)	0.56 (*1.03*)	0.50 (*.89*)
	95% CI	[-0.16, 0.66]	[-0.15, 0.78]	[-0.14, 0.39]	[-0.06, 0.56]	[0.01, 1.11]	[0.02, 0.98]
	Min-Max	0.00-3.00	0.00-3.00	0.00-2.00	0.00-2.00	0.00-3.00	0.00-3.00
Intensive (17)	Mean (*SD*)	0.35 (*.49*)	5.53 (*2.45*)	5.00 (*2.37*)	0.76 (*1.09*)	5.06 (*2.44*)	4.47 (*2.65*)
	95% CI	[0.10, 0.61]	[4.27, 6.79]	[3.78, 6.22]	[0.20, 1.33]	[3.81, 6.31]	[3.11, 5.83]
	Min-Max	0.00-1.00	1.00-8.00	1.00-8.00	0.00-3.00	1.00-9.00	0.00-9.00
Spaced (14)	Mean (*SD*)	0.21 (*.58*)	5.36 (*2.27*)	5.36 (*2.37*)	0.29 (*.47*)	5.36 (*2.27*)	4.07 (*2.40*)
	95% CI	[-0.12, 0.55]	[4.04, 6.67]	[3.99, 6.73]	[0.02, 0.56]	[4.04, 6.67]	[2.69, 5.46]
	Min-Max	0.00-2.00	1.00-8.00	1.00-8.00	0.00-1.00	1.00-8.00	0.00-9.00

The descriptive statistics in Tables 4.23 and 4.24 show that the students in both Experimental Groups (Intensive and Spaced) made very great progress in their Passive and Active Knowledge, as demonstrated in the immediate post-test for both the Recognition and Recall categories. Moreover, this improvement was sustained in the delayed post-test. However, the descriptive statistics indicate that both Experimental Groups had slightly greater improvement in Recognition than Recall questions. Conversely, the Control Group's mean scores over the three time points (pre-, immediate post-, and delayed post-tests) reflected no change in the students' Passive or Active Knowledge, whether in the Recognition or Recall categories.

4.5.4.1 Passive and Active Recognition for the 4-day RI Sub-groups

The students' scores for the Passive and Active Recognition items at all three time points (pre-, immediate post-, and delayed post-) for the 4-day RI sub-groups were then submitted to two-way mixed ANOVA. The results for Passive Recognition subsequently revealed a

significant main effect of Time: $F(1.7, 88)=147.25$, $p< .001$, $\eta_{p2}= .770$, $\eta_2= .576$. The main Group effect was also statistically significant: $F(2, 44)=38.12$, $p< .001$, $\eta_{p2}= .634$, $\eta_2= .328$, as was the Group * Time interaction: $F(3.48, 88)=32.29$, $p< .001$, $\eta_{p2}= .595$, $\eta_2= .252$. Similarly, the Active Recognition results showed a significant main effect of Time: $F(1.5, 88)=119.81$, $p< .001$, $\eta_{p2}= .731$, $\eta_2= .576$, and the main Group effect was likewise statistically significant, $F(2, 44)=15.35$, $p< .001$, $\eta_{p2}= .411$, $\eta_2= .353$, as was the Group * Time interaction: $F(2.91, 88)=22.16$, $p< .001$, $\eta_{p2}= .502$, $\eta_2= .213$.

Table 4.25

Effect Sizes for Repeated-measures (within-subject) Comparisons between the Different Test Time Points for Each Group (4-day RI Sub-groups)

Group		Passive Recognition			Active Recognition	
		Time 2	Time 3		Time 2	Time 3
Control	**Time 1**	*p*=1.00, *d*=0.00, 95% CI [-0.98, 0.98]	*p*= .685, *d*=0.33, 95% CI [-0.66, 1.31]	**Time 1**	*p*=1.00, *d*=0.08, 95% CI [-0.91, 1.06]	*p*= .217, *d*=0.36, 95% CI [-0.63, 1.35]
	Time 2		*p=.290*, *d*=0.35, 95% CI [-0.64, 1.34]	**Time 2**		*p*< .033, *d*=0.26, 95% CI [-0.72, 1.25]
Intensive	**Time 1**	*p*< .001, *d*=3.39, 95% CI [1.91, 4.88]	*p*< .001, *d*=2.89, 95% CI [1.53, 4.25]	**Time 1**	*p*< .001, *d*=2.56, 95% CI [1.28, 3.85]	*p*< .001, *d*=2.10, 95% CI [0.91, 3.28]
	Time 2		*p*=1.00, *d*= -0.03, 95% CI [-0.98,0.92]	**Time 2**		*p*= .499, *d*= -0.22, 95% CI [-1.17, 0.74]
Spaced	**Time 1**	*p*< .001, *d*=4.60, 95% CI [2.60, 6.61]	*p*< .001, *d*=4.21, 95% CI [2.33, 6.09]	**Time 1**	*p*< .001, *d*=2.75, 95% CI [1.29, 4.21]	*p*< .001, *d*=2.62, 95% CI [1.19, 4.04]
	Time 2		*p*=1.00, *d*= -0.21, 95% CI [-1.28, 0.83]	**Time 2**		*p*=1.00, *d*= -0.15, 95% CI [-1.20, 0.90]

Table 4.26

Between-subject, Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (4-day Sub-groups)

Group		Passive Recognition			Active Recognition	
		Intensive	Spaced		Intensive	Spaced
Time 1	**Control**	p=1.00, d=0.00, 95% CI [-0.68, 0.68]	p=1.00, d=-0.04, 95% CI [-0.76, 0.68]	**Control**	p=1.00, d=0.58, 95% CI [-0.12, 1.28]	p=1.00, d=0.08, 95% CI [-0.64, 0.79]
	Intensive		p=1.00, d=0.05, 95% CI [-0.76, 0.66]	**Intensive**		p=1.00, d=-0.54, 95% CI [-1.26, 0.18]
Time 2	**Control**	p< .001, d=3.04, 95% CI [2.03, 4.04]	p< .001, d=4.18, 95% CI [2.90, 5.46]	**Control**	p< .001, d=2.38, 95% CI [1.49, 3.27]	p< .001, d=2.79, 95% CI [1.78, 3.80]
	Intensive		p=1.00, d=0.33, 95% CI [-0.39, 1.04]	**Intensive**		p=1.00, d= -0.13, 95% CI [-0.58, 0.84]
Time 3	**Control**	p< .001, d=2.42, 95% CI [1.52, 3.32]	p< .001, d=3.66, 95% CI [2.49, 4.83]	**Control**	p< .001, d=1.98, 95% CI [1.15, 2.82]	p< .001, d=2.03, 95% CI [1.15, 2.91]
	Intensive		p=1.00, d= -0.18, 95% CI [-0.53, 0.89]	**Intensive**		p=1.00, d= -0.16, 95% CI [-0.87, 0.55]

In Table 4.25, the results show that significant progress was made by the Intensive and Spaced Groups in both the Passive and Active Recognition sub-tests, between Time 1 and Time 2, and between Time 1 and Time 3. This progress is indicated by the very large effect sizes (see Table 4.25), with confidence intervals that do not cross zero, meaning that the effect sizes are reliable. However, the Intensive and Spaced Groups did not make any significant progress between Time 2 and Time 3. This indicates that the Experimental Groups upheld their higher scores in the Passive and Active Recognition sub-tests at the delayed post-test time point. Although all the Experimental Group effect sizes are very large, the largest gains were for passive recognition, as reflected in the even larger effect sizes than for active recognition. This demonstrates that L2-L1 recognition (Passive) may have been easier for the participants than the recognition of the L1-L2 (Active).

Conversely, no significant differences were demonstrated by the Control Group between Time 1 and Time 2, or between Time 1 and Time 3 in either the Passive or Active Recognition sub-tests. However, significant change was revealed amongst the Control Group in the Active Recognition sub-test between Times 2 and 3. Meanwhile, the Control Group's scores at Time 3 remained lower than those of the two Experimental Groups (Intensive and Spaced) (see Table 4.26 for these results in detail). However, the effect size of the comparison between Times 2 and 3 was very small for the Control Group, with the confidence interval crossing zero. This slight increase in the score results of the Control Group may be due to the fact that there were only four days between the immediate and delayed post-tests, and this increase might, therefore, reflect a small test effect (see Table 4.25). Generally, there were no differences between the Experimental Groups' scores for either Passive or Active Recognition over the three time points (see Figures 4.21 and 4.22).

Figure 4.21
Mean Scores for the Passive Recognition Vocabulary Task over Time and by Group (4-day RI Sub-groups)

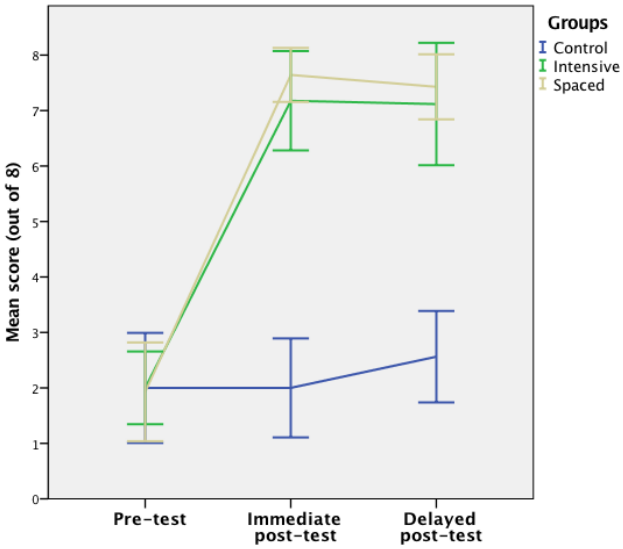
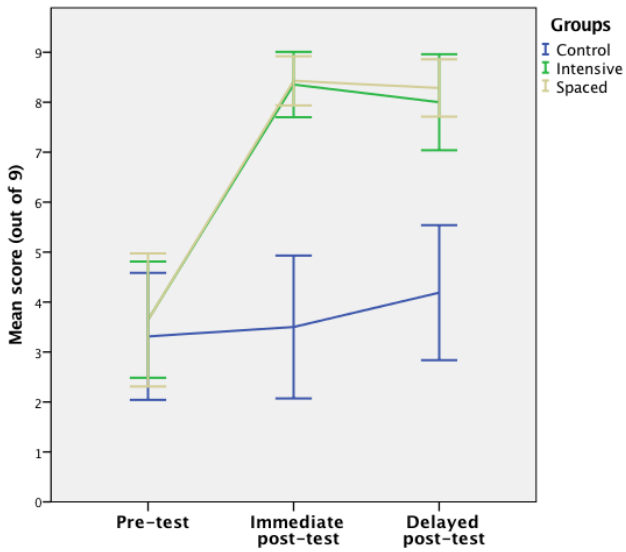


Figure 4.22
Mean Scores for the Active Recognition Vocabulary Task over Time and by Group (4-day RI Sub-groups)



The results for Passive Recall indicated that there was a significant main effect of Time: $F(1.4, 88)=143.38$, $p< .001$, $\eta_{p2}= .765$, $\eta_2= .548$. The main Group effect was also significant: $F(2, 44)=32.05$, $p< .001$, $\eta_{p2}= .593$, $\eta_2= .366$. Likewise, the Group * Time interaction was statistically significant: $F(2.74, 88)=37.13$, $p< .001$, $\eta_{p2}= .628$, $\eta_2= .284$. Correspondingly, the results for Active Recall showed significant main effects of Time: $F(1.2, 88)=85.28$, $p< .001$, $\eta_{p2}= .660$, $\eta_2= .518$, and the main Group effect: $F(2, 44)=16.99$, $p< .001$, $\eta_{p2}= .436$, $\eta_2= .355$; as well as the Group * Time interaction: $F(2.49, 88)=17.69$, $p< .001$, $\eta_{p2}= .446$, $\eta_2= .215$.

Table 4.27

Effect Sizes for Repeated Measures (within-subject) Comparisons between the Different Test Time Points for Each Group (4-day RI Sub-groups)

Group		Passive Recall			Active Recall	
		Time 2	Time 3		Time 2	Time 3
Control	Time 1	*p*=1.00, *d*=0.07, 95% CI [-0.91, 1.05]	*p*=1.00, *d*=0.09, 95% CI [-0.89, 1.07]	Time 1	*p*=1.00, *d*=0.37, 95% CI [-0.62,1.36]	*p*=1.00, *d*=0.33, 95% CI [-0.65, 1.32]
	Time 2		*p*=1.00, *d*=0.00, 95% CI [-0.98, 0.98]	Time 2		*p*=1.00, *d*=-0.06, 95% CI [-1.04, 0.92]
Intensive	Time 1	*p*< .001, *d*=2.93, 95% CI [1.56, 4.30]	*p*< .001, *d*=2.72, 95% CI [1.40, 4.04]	Time 1	*p*< .001, *d*=2.28, 95% CI [1.06, 3.50]	*p*< .001, *d*=1.83, 95% CI [0.70, 2.96]
	Time 2		*p*=.054, *d*= -0.22, 95% CI [-1.17, 0.73]	Time 2		*p*<.009, *d*= -0.23, 95% CI [-1.19, 0.72]
Spaced	Time 1	*p*< .001, *d*=3.11, 95% CI [1.55, 4.67]	*p*< .001, *d*=2.99, 95% CI [1.46, 4.51]	Time 1	*p*< .001, *d*=3.09, 95% CI [1.54, 4.65]	*p*< .001, *d*=2.19, 95% CI [0.86, 3.51]
	Time 2		*p*=1.00, *d*=0.00, 95% CI [-1.05, 1.05]	Time 2		*p*=1.00, *d*= -0.55, 95% CI [-1.62, 0.52]

Table 4.28

Between-subject, Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (4-day Sub-groups)

Group		Passive Recall			Active Recall	
		Intensive	Spaced		Intensive	Spaced
	Control	*p*=1.00, *d*=0.16, 95% CI [-0.53, 0.84]	*p*=1.00, *d*= -0.06, 95% CI [-0.78, 0.66]	**Control**	*p*= .196, *d*=0.58, 95% CI [-0.12,1.28]	*p*=1.00, *d*=0.08, 95% CI [-0.64, 0.79]
Time 1						
	Intensive		*p*=1.00, *d*= -0.26, 95% CI [-0.97, 0.45]	**Intensive**		*p*= .290, *d*= -0.54, 95% CI [-1.26, 0.18]
	Control	*p*< .001, *d*=2.81, 95% CI [1.84, 3.77]	*p*< .001, *d*=3.02, 95% CI [1.97, 4.07]	**Control**	*p*< .001, *d*=2.38, 95% CI [1.49, 3.27]	*p*< .001, *d*=2.79, 95% CI [1.78, 3.80]
Time 2						
	Intensive		*p*=1.00, *d*= -0.07, 95% CI [-0.78, 0.64]	**Intensive**		*p*= .743, *d*= -0.13, 95% CI [-0.58, 0.84]
	Control	*p*< .001, *d*=2.80, 95% CI [1.84, 3.76]	*p*< .001, *d*=3.16, 95% CI [2.09, 4.23]	**Control**	*p*< .001, *d*=1.98, 95% CI [1.15, 2.82]	*p*< .001, *d*=2.03, 95% CI [1.15, 2.91]
Time 3						
	Intensive		*p*=1.00, *d*=0.15, 95% CI [-0.56, 0.86]	**Intensive**		*p*=1.00, *d*= -0.16, 95% CI [-0.87, 0.55]

Figure 4.23

Mean Scores for the Passive Recall Vocabulary Task over Time and by Group (4-day RI Sub-groups)

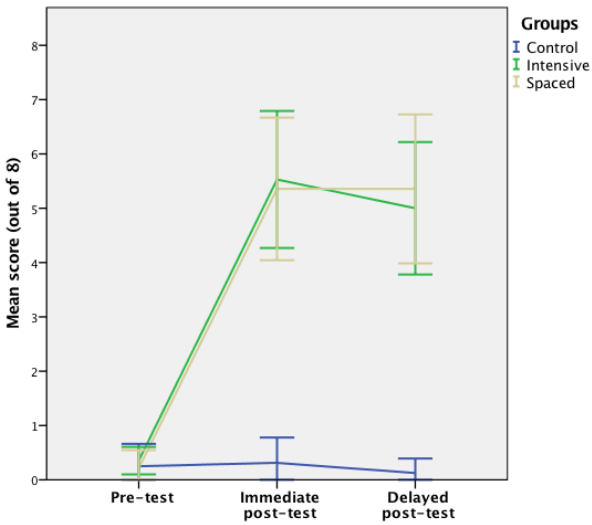
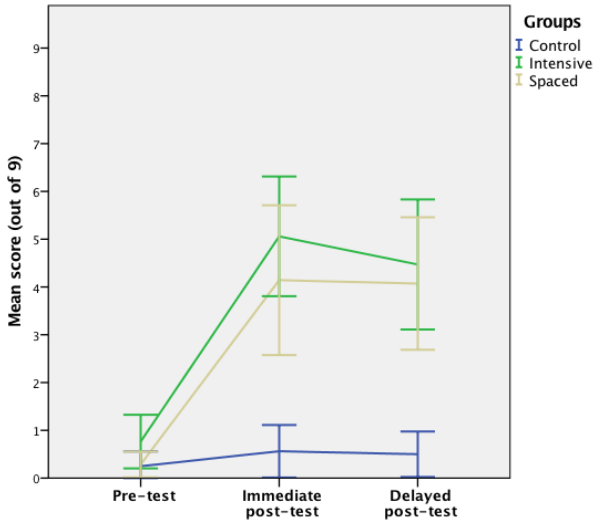


Figure 4.24

Mean Scores for the Active Recall Vocabulary Task over Time and by Group (4-day RI Sub-groups)



Meanwhile, the results of comparisons between the groups, illustrated in Table 4.28, and Figures 4.23 and 4.24, reveal that the Experimental Groups produced equally satisfactory results, with no short- or long-term differences in either the Passive or Active Recall sub-tests at the immediate and delayed (4-day RI) post-test time points . Although the Intensive Group's scores dropped significantly for the Active Recall sub-test between Time 2 and Time 3, the very small effect sizes (d= -0.23) and confidence intervals crossing zero indicate that this effect was unreliable (Plonsky & Oswald, 2014). In addition, Table 4.27 shows that there were no significant differences between the scores of either the Intensive and Spaced Groups in the Active and Passive Recall sub-tests. However, the Control Group demonstrated no improvement in their scores for the Passive and Active sub-tests over the three time points. Moreover, the Control Group's performance was significantly lower than that of the Intensive and Spaced Groups in both the immediate and delayed post-tests, which is reflected in the large between-group effect sizes (see Table 4.28).

4.5.5 The 28-day RI Sub-groups: Passive and Active Vocabulary Knowledge

This section focuses on the results of the two sub-tests (Passive and Active) in each main category, Recognition and Recall, for the 28-day RI sub-groups (see descriptive statistical data in Table 4.29 for Passive/Active Recognition Knowledge and Table 4.30 for Passive/Active Recall Knowledge).

Table 4.29

Descriptive Statistical Data for Recognition Knowledge (Passive/ Active) at Each Test Time Point (28-day RI sub-groups)

Group (N)		Recognition Knowledge					
		Passive			Active		
		Pre-test	Immediate Post-test	Delayed Post-test (28-day RI)	Pre-test	Immediate Post-test	Delayed Post-test (28-day RI)
Control (17)	Mean (*SD*)	2.71 (*1.69*)	2.41 (*1.87*)	2.65 (*2.06*)	4.00 (*3.08*)	4.47 (*2.76*)	4.24 (*2.61*)
	95% CI	[1.84, 3.57]	[1.45, 3.37]	[1.59, 3.71]	[2.42, 5.58]	[3.05, 5.89]	[2.89, 5.58]
	Min-Max	0.00-6.00	0.00-6.00	0.00-7.00	0.00-9.00	0.00-9.00	0.00-9.00
Intensive (17)	Mean (*SD*)	1.65 (*1.50*)	7.71 (*0.47*)	7.18 (*1.07*)	3.71 (*2.47*)	8.35 (*1.22*)	7.94 (*1.64*)
	95% CI	[0.88, 2.42]	[7.46, 7.95]	[6.62, 7.73]	[2.44, 4.98]	[7.72, 8.98]	[7.10, 8.78]
	Min-Max	0.00-5.00	7.00-8.00	5.00-8.00	0.00-8.00	5.00-9.00	5.00-9.00
Spaced (15)	Mean (*SD*)	2.47 (*2.17*)	7.20 (2.*04*)	7.07 (*1.67*)	3.60 (*2.50*)	8.07 (*1.83*)	7.87 (*1.73*)
	95% CI	[1.27, 3.67]	[6.07, 8.33]	[6.14, 7.99]	[2.21, 4.99]	[7.05, 9.08]	[6.91, 8.82]
	Min-Max	0.00-6.00	0.00-8.00	3.00-8.00	0.00-8.00	2.00-9.00	3.00-9.00

Table 4.30

Descriptive Statistical Data for Recall Knowledge (Passive/ Active) at Each Test Time Point (28-day RI Sub-groups)

Group (N)		Recall Knowledge					
		Passive			Active		
		Pre-test	Immediate Post-test	Delayed Post-test (28-day RI)	Pre-test	Immediate Post-test	Delayed Post-test (28-day RI)
Control (17)	Mean (*SD*)	0.35 (*1.00*)	0.41 (*.87*)	0.59 (*1.28*)	0.41 (*.71*)	1.12 (*2.18*)	0.47 (*.80*)
	95% CI	[-0.16, 0.87]	[-0.04, 0.86]	[-0.07, 1.25]	[0.05, 0.78]	[0.00, 2.24]	[0.06, 0.88]
	Min-Max	0.00-4.00	0.00-3.00	0.00-5.00	0.00-2.00	0.00-9.00	0.00-2.00
Intensive (17)	Mean (*SD*)	0.29 (*.59*)	4.94 (*3.01*)	3.29 (*2.89*)	0.29 (*.47*)	4.00 (*3.37*)	3.12 (*2.50*)
	95% CI	[-0.01, 0.60]	[3.39, 6.49]	[1.81, 4.78]	[0.05, 0.54]	[2.27, 5.73]	[1.83, 4.40]
	Min-Max	0.00-2.00	0.00-8.00	0.00-8.00	0.00-1.00	0.00-9.00	0.00-8.00
Spaced (15)	Mean (*SD*)	0.33 (*.62*)	5.20 (*2.83*)	4.07 (*3.01*)	0.53 (*.92*)	3.93 (*2.69*)	3.20 (*2.37*)
	95% CI	[-0.01, 0.68]	[3.63, 6.77]	[2.40, 5.73]	[0.03, 1.04]	[2.45, 5.42]	[1.89, 4.51]
	Min-Max	0.00-2.00	0.00-8.00	0.00-8.00	0.00-2.00	0.00-9.00	0.00-8.00

The descriptive statistics in Tables 4.29 and 4.30 illustrate that the students in the Intensive and Spaced Groups made remarkable progress in their Passive and Active Knowledge for both Recognition and Recall in the immediate post-test. This progress was maintained in the Recognition category of the delayed post-test, but dropped in the Recall

category of the delayed post-test. However, the Control Group showed no differences in their scores for the Passive and Active sub-tests over the three tests (pre-, immediate post-, and delayed post-tests).

4.5.5.1 Passive and Active Recognition for the 28-day RI Sub-groups

The Passive and Active Recognition results achieved by the 28-day RI sub-groups in the pre-, immediate post-, and delayed post-tests, obtained by running two-way mixed ANOVAs, revealed a significant main effect of Time for Passive Recognition: $F(1.87, 92)=128.55$, $p< .001$, $\eta_{p2}= .736$, $\eta_2= .512$. Additionally, the main Group effect was significant: $F(2, 46)= 26.38$, $p< .001$, $\eta_{p2}= .534$, $\eta_2= .280$. The Group * Time interaction was also statistically significant: $F(3.74, 92)=38.33$, $p< .001$, $\eta_{p2}= .625$, $\eta_2= .305$. This significant interaction is evident from the very large effect sizes for the two Experimental Groups compared with the very small effect sizes for the Control Group (Table 4.31). Correspondingly, the results for Active Recognition indicated a similarly significant main effect of Time: $F(1.89, 92)=79.17$, $p< .001$, $\eta_{p2}= .633$, $\eta_2= .505$. In addition, the main Group effect was also statistically significant: $F(2, 46)=7.82$, $p< .001$, $\eta_{p2}= .254$, $\eta_2= .235$, as was the Group * Time interaction: $F(3.77, 92)=15.79$, $p< .001$, $\eta_{p2}= .407$, $\eta_2= .201$. This statistically significant difference among the groups over the three time points is reflected in the corresponding large effect sizes (see section 3.5.2).

Table 4.31

Effect Sizes for Repeated Measures (within-subject) Comparisons between the Different Test Time Points for Each Group (28-day RI sub-groups)

		Passive Recognition			Active Recognition	
Group		**Time 2**	**Time 3**		**Time 2**	**Time 3**
Control	**Time 1**	*p*=1.00, *d*= -0.17, 95% CI [-1.12, 0.78]	*p*=1.00, *d*= -0.03, 95% CI [-0.98, 0.92]	**Time 1**	*p*= .892, *d*=0.16, 95% CI [-0.79, 1.11]	*p*=1.00, *d*=0.08, 95% CI [-0.87, 1.04]
	Time 2		*p*=1.00, *d*=0.12, 95% CI [-0.83, 1.07]	**Time 2**		*p*=1.00, *d*= -0.09, 95% CI [-1.04, 0.87]
Intensive	**Time 1**	*p*< .001, *d*=5.45, 95% CI [3.39, 7.52]	*p*< .001, *d*=4.25, 95% CI [2.53, 5.96]	**Time 1**	*p*< .001, *d*=2.38, 95% CI [1.14, 3.63]	*p*< .001, *d*=2.02, 95% CI [0.85, 3.19]
	Time 2		*p*= .478, *d*= -0.64, 95% CI [-1.62, 0.33]	**Time 2**		*p*=1.00, *d*= -0.28, 95% CI [-1.24, 0.67]
Spaced	**Time 1**	*p*< .001, *d*=2.25, 95% CI [0.95, 3.54]	*p*< .001, *d*=2.38, 95% CI [1.05, 3.70]	**Time 1**	*p*< .001, *d*=2.04, 95% CI [0.79, 3.29]	*p*< .001, *d*=1.99, 95% CI [0.75, 3.22]
	Time 2		*p*=1.00, *d*= -0.07, 95% CI [-1.08, 0.94]	**Time 2**		*p*=1.00, *d*= -0.11, 95% CI [-1.12, 0.90]

Table 4.32

Between-subject, Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (28-day Sub-groups)

		Passive Recognition			Active Recognition	
Group		**Intensive**	**Spaced**		**Intensive**	**Spaced**
Time 1	**Control**	*p*= .273, *d*= -0.66, 95% CI [-1.37, 0.04]	*p*=1.00, *d*= -0.12, 95% CI [-0.82, 0.57]	**Control**	*p*=1.00, *d*= -0.10, 95% CI [-1.06, 0.85]	*p*=1.00, *d*= -0.14, 95% CI [-0.84, 0.55]
	Intensive		*p*= .607, *d*=0.45, 95% CI [-0.26, 1.15]	**Intensive**		*p*=1.00, *d*= -0.04, 95% CI [-0.74, 0.65]
Time 2	**Control**	*p*< .001, *d*=3.89, 95% CI [2.27, 5.50]	*p*< .001, *d*=2.46, 95% CI [1.54, 3.37]	**Control**	*p*< .001, *d*=1.82, 95% CI [0.69, 2.95]	*p*< .001, *d*= -0.18, 95% CI [-0.88, 0.51]
	Intensive		*p*=1.00, *d*= -0.36, 95% CI [-1.06, 0.34]	**Intensive**		*p*=1.00, *d*=1.52, 95% CI [0.73, 2.31]
Time 3	**Control**	*p*< .001, *d*=2.76, 95% CI [1.43, 4.09]	*p*< .001, *d*=2.34, 95% CI [1.44, 3.24]	**Control**	*p*< .001, *d*=1.70, 95% CI [0.59, 2.81]	*p*< .001, *d*=1.62, 95% CI [0.82, 2.42]
	Intensive		*p*=1.00, *d*= -0.08, 95% CI [-0.77, 0.62]	**Intensive**		*p*=1.00, *d*= -0.04, 95% CI [-0.74, 0.65]

Figure 4.25
Mean Scores for the Passive Recognition Vocabulary Task over Time and by Group (28-day RI Sub-groups)

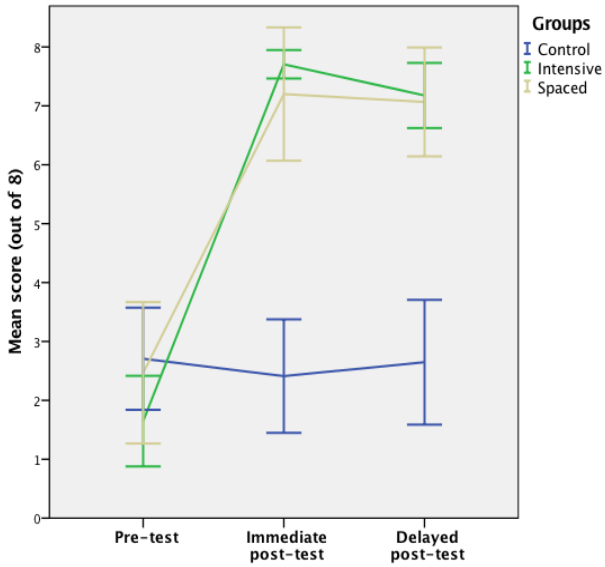
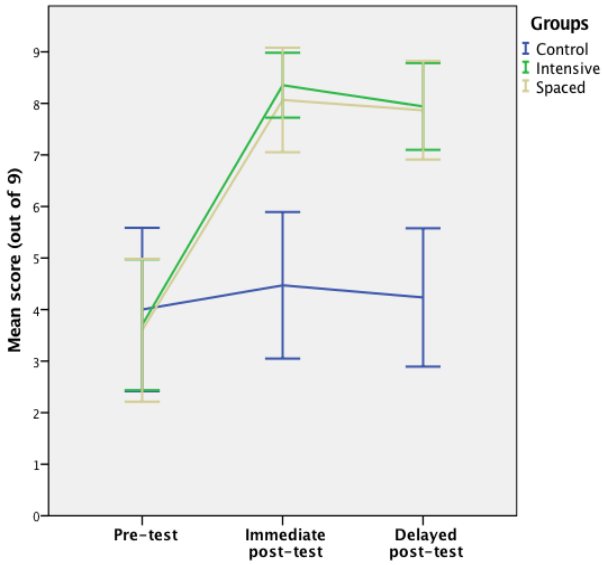


Figure 4.26
Mean Scores for the Active Recognition Vocabulary Task over Time and by Group (28-day RI Sub-groups)



The results indicate that significant progress was made by both the Intensive and Spaced Groups in the Passive and Active Recognition sub-tests between Times 1 and 2, and between Times 1 and 3. This progress is evident from the very large effect sizes (see Table 4.31), with confidence intervals that do not cross zero, meaning that the effect sizes are reliable. Conversely, the Control Group did not make any significant progress at any of the time points and this is supported by a small effect size and confidence intervals that cross zero. Furthermore, there were no significant differences demonstrated by the Experimental Groups (Intensive and Spaced) between Times 2 and 3, suggesting that they maintained their higher scores in the delayed post-test. However, the Control Group's scores at Time 3 remained lower than those of the two Experimental Groups. Table 4.32 presents detailed comparisons between the groups' results. Although gains for Passive Recognition were slightly higher than for Active Recognition, generally there were no differences between the results for either Passive or Active Recognition, as shown in Figures 4.25 and 4.26.

4.5.5.2 Passive and Active Recall for the 28-day RI Sub-groups

The scores for Passive Recall displayed a main effect of Time: $F(1.56, 92)=66.86$, $p< .001$, $\eta_{p2}= .592$, $\eta_2= .460$. The main Group effect was likewise statistically significant: $F(2, 46)=12.72$, $p< .001$, $\eta_{p2}= .356$, $\eta_2= .277$, as was the Group * Time interaction: $F(3.12, 92)= 16.19$, $p< .001$, $\eta_{p2}= .413$, $\eta_2= .223$. The large effect sizes indicate this significant interaction between the Group and time for the Experimental Groups, with small effect sizes for the Control Group. Similarly, the Active Recall data revealed a significant main effect of Time: $F(1.50, 92)= 43.67$, $p< .001$, $\eta_{p2}= .487$, $\eta_2= .420$. The main Group effect was also statistically significant: $F(2, 46)=6.76$, $p< .003$, $\eta_{p2}= .227$, $\eta_2= .214$, as was the Group * Time interaction: $F(3.00, 92)=7.16$, $p< .001$, $\eta_{p2}= .237$, $\eta_2= .138$.

Table 4.33

Effect Sizes for Repeated Measures (within-subject) Comparisons between the Different Test Time Points for Each Group (28-day RI Sub-groups)

		Passive Recall			Active Recall	
Group		**Time 2**	**Time 3**		**Time 2**	**Time 3**
Control	**Time 1**	*p*=1.00, *d*=0.06, 95% CI [-0.89, 1.02]	*p*=1.00, *d*=0.21, 95% CI [-0.74, 1.16]	**Time 1**	*p*=759, *d*=0.44, 95% CI [-0.52,1.4]	*p*=1.00, *d*=0.08, 95% CI [-0.87, 1.03]
	Time 2		*p*=1.00, *d*=0.16, 95% CI [-0.79, 1.12]	**Time 2**		*p*= .401, *d*= -0.40, 95% CI [-1.36, 0.56]
Intensive	**Time 1**	*p*< .001, *d*=2.14, 95% CI [0.95, 3.34]	*p*< .001, *d*=1.44, 95% CI [0.37, 2.51]	**Time 1**	*p*< .001, *d*=1.54, 95% CI [0.46, 2.63]	*p*< .001, *d*=1.57, 95% CI [0.49, 2.66]
	Time 2		*p*<001, *d*= -0.56, 95% CI [-1.53, 0.41]	**Time 2**		*p*= .129, *d*= -0.30, 95% CI [-1.25, 0.66]
Spaced	**Time 1**	*p*< .001, *d*= 2.38, 95% CI [1.06, 3.62]	*p*< .001, *d*=1.72, 95% CI [0.54, 2.91]	**Time 1**	*p*< .001, *d*= 1.69, 95% CI [0.58, 2.80]	*p*< .001, *d*=1.49, 95% CI [0.41, 2.56]
	Time 2		*p*< .007, *d*= -0.39, 95% CI [-1.41, 0.64]	**Time 2**		*p*= .333, *d*= -0.29, 95% CI [-1.24, 0.67]

Table 4.34

Between-subject, Post Hoc Comparisons between Groups (Control, Intensive and Spaced) at Each Time Point (28-day Sub-groups)

Group		Passive Recall			Active Recall	
		Intensive	Spaced		Intensive	Spaced
Time 1	**Control**	p=1.00, d=1.02, 95% CI [0.01, 2.03]	p=1.00, d= -0.02, 95% CI [-0.72, 0.67]	**Control**	p=1.00, d= -0.20, 95% CI [-1.15, 0.75]	p=1.00, d=0.15, 95% CI [-0.55, 0.84]
	Intensive		p=1.00, d=-0.07, 95% CI [-0.63, 0.76]	**Intensive**		p=1.00, d= -0.34, 95% CI [-0.36, 1.03]
Time 2	**Control**	p< .001, d=2.05, 95% CI [0.87, 3.22]	p< .001, d=2.35, 95% CI [1.45, 3.26]	**Control**	p< .013, d=1.02, 95% CI [0.01, 2.03]	p< .020, d=1.16, 95% CI [0.41, 1.91]
	Intensive		p=1.00, d=-0.09, 95% CI [-0.61, 0.78]	**Intensive**		p=1.00, d= -0.02, 95% CI [-0.72, 0.67]
Time 3	**Control**	p< .008, d=1.21, 95% CI [0.17, 2.24]	p< .001, d=1.54, 95% CI [0.75, 2.33]	**Control**	p< .001, d=1.43, 95% CI [0.36, 2.49]	p< .001, d=1.59, 95% CI [0.79, 2.38]
	Intensive		p=1.00, d=0.27, 95% CI [-0.43, 0.96]	**Intensive**		p=1.00, d=0.03, 95% CI [-0.66, 0.73]

Figure 4.27

Mean Scores for the Passive Recall Vocabulary Task over Time and by Group (28-day RI Sub-groups).

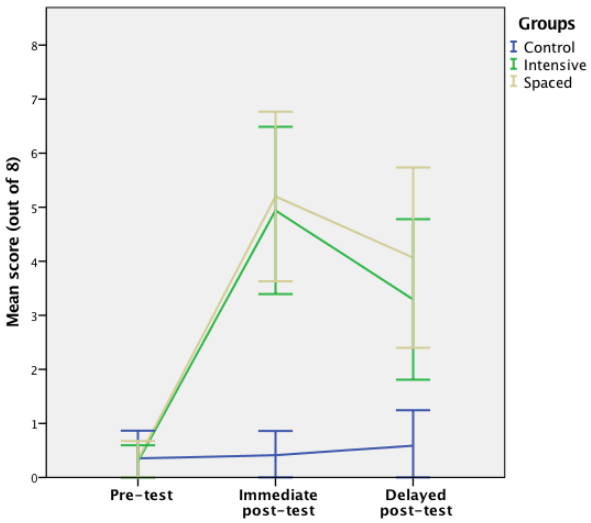
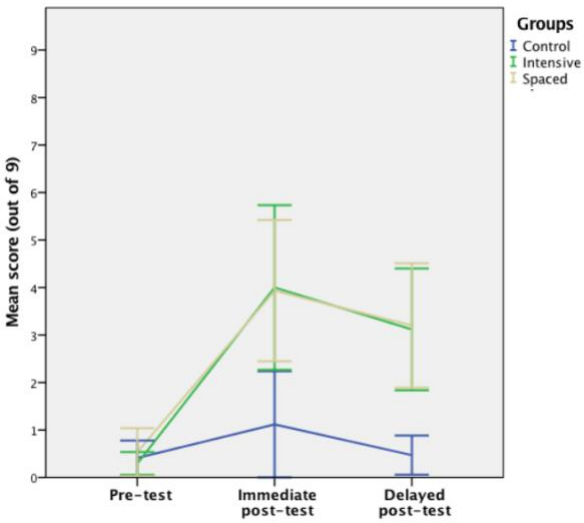


Figure 4.28

Mean Scores for the Active Recall Vocabulary Task over Time and by Group (28-day RI Sub-groups).



Similarly, the results of the Experimental Groups, illustrated in Tables 4.33 and Figures 4.27 and 4.28, reveal that the Intensive and Spaced Groups both made significant progress in the immediate post-test, with very large and reliable effect sizes (see Table 4.33). However, the scores from both Experimental Groups dropped in the Passive and Active sub-tests at the delayed (28-day RI) post-test time point, with no statistically significantly differences. In addition, the Experimental Groups significantly outperformed the Control Group in the immediate and delayed (28-day RI) post-tests, and this is reflected in the corresponding large effect sizes (see Table 4.34). Conversely, the Control Group did not improve across the three different time points, whether in the Passive or Active sub-tests, as is clearly shown in Table 4.33. Moreover, there were no significant differences in Active Recall amongst the Experimental Groups (Intensive and Spaced) at the delayed (28-day RI) post-test time point, and between Times 2 and 3, which suggests that the Experimental Groups maintained their scores (Figure 4.28). Meanwhile there was a significant falling-off in performance among the Intensive and Spaced Groups in their Passive Recall between Times 2 and 3 (see Figure 4.27). However, the effect sizes here were small, and the confidence intervals crossed zero, suggesting that these effects were unreliable (see Table 4.33).

4.6 Summary

To summarise, the findings for experimental learning gains were used to answer the first and second research question in this study. It was interesting that no significant differences were found between the acquisition (immediate post-test) and retention (delayed post-test) of target vocabulary words by the Intensive and Spaced Groups. At a descriptive level, the Spaced Group lost less vocabulary knowledge than the Intensive Group in terms of Recall, although the effect size for the Intensive Group was small, meaning that it does not represent a reliable difference between the two groups. Regarding short-term and long-term vocabulary acquisition, the results for Recognition in the delayed post-test showed that the Experimental Groups maintained their progress in terms of vocabulary learning gains. However, the Recall results dropped statistically significantly at delayed post-test. These quantitative results also revealed that both the Intensive and Spaced Groups significantly outperformed the Control Group in their short- and long-term vocabulary learning.

Additionally, the quantitative results indicated that there were no significant differences between the two sub-groups (4-day RI; 28-day RI), which contradicts the previous literature (for example, Bird, 2010; Rogers, 2015), where spaced groups have been shown to retain more information by the delayed post-test (see Discussion Chapter for more details of this). When the results of each Recognition and Recall were analysed, based on the Passive and Active sub-tests, the scores for Passive and Active Recognition showed some similarity to the Passive and Active Recall scores over the time points for the 4-day RI and 28-day RI sub-groups. The results of the Active Recall sub-tests for the Intensive Group at the delayed (4-day RI) post-test time point, as well as the results of the Passive Recall sub-tests for the 28-day sub-groups (Intensive and Spaced), showed a significant decline from Time 2 to Time 3. However, these results are not reliable, because of the small effect sizes and confidence intervals crossing zero. Generally, it is noted that the gains for Recognition are higher than for Recall, whether it is Passive or Active. Overall, however, the results suggest that the use of Quizlet helped the learners considerably with their vocabulary learning, which is in line with previous findings in the literature (for example, Dizon, 2016; see Discussion Chapter).

The next chapter will focus on the perceptions of teachers and students with regard to using Quizlet in their vocabulary learning. These findings are aimed at answering the first and third research questions in this study.

CHAPTER FIVE: RESULTS 2 – STUDENTS' AND TEACHERS' PERCEPTIONS OF QUIZLET

5.1 Introduction

The key focus of this current chapter is to examine the participants' (students' and teachers') perceptions of using the Quizlet software programme for vocabulary learning, corresponding to the first and the third research questions (see section 4.2). Therefore, both quantitative (survey/questionnaire) and qualitative (open-ended questionnaire items and interview) data were collected from both sets of participants. All the data were organised under three main themes: perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI) (Davis, 1989). This chapter will begin with the questionnaire results, in order to discover the students' views relating to the first research question. The chapter will then move on to the teachers' qualitative (interview) data, so that the third research question can be addressed. This will be followed by a summary of the participants' perceptions findings.

5.2 Quizlet Questionnaire: The Students' Perceptions

A questionnaire (see subsection 3.4.4.2) was administered to the students, in order to gather their views on using Quizlet. It should be noted that only responses from students who had attended at least three training sessions were added to the Quizlet questionnaire data. Overall, seven (of 94) students were excluded: two from the Intensive Group and five from the Spaced Group, because they had attended just two sessions. This decision to include students who had only attended three out of the four sessions was due to the fact that this represented 75% of the training. Thus, it was anticipated these students would have sufficient knowledge of the Quizlet software programme to be able to complete the questionnaire.

Before conducting the analysis, the normality and homogeneity of variance was checked to determine whether parametric (an independent samples t-test) or non-parametric (Kruskal-Wallis or Mann-Whitney) tests should be used to identify any significant differences in the groups' mean scores from the questionnaires. Therefore, the Shapiro-Wilk and Levene's tests ($p > .05$) were applied to the three themes (PEOU, PU and BI), as the analysis would be based on these (see Table 5.1).

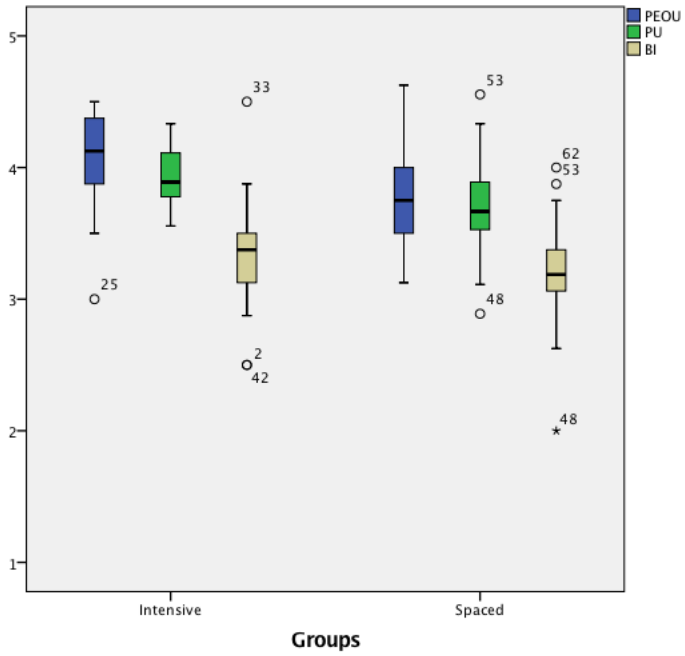
Table 5.1
Shapiro-Wilk Test of Normality and Homogeneity of Variance for the Themes, PEOU, PU and BI

Theme		**Shapiro-Wilk Test**		**Levene's Test**
		df	*P*-value	*P*-value
PEOU	Intensive	47	.010	.576
	Spaced	40	.273	
PU	Intensive	47	.065	.019
	Spaced	40	.831	
BI	Intensive	47	.002	.808
	Spaced	40	.007	

Table 5.1 shows that the assumptions of homogeneity of variance for the themes, PEOU and BI, were satisfied, using the Levene's test ($p > .05$). However, the PU theme was found to violate the assumption of homogeneity of variance, using the same test ($p < .05$). Moreover, the assumption of normality for the three themes, addressed using the Shapiro-Wilk test, produced results that differed between the Intensive and Spaced Groups. Thus, the results for PEOU in the Spaced Group met the assumption of normality, but the Intensive Group's results violated the assumption for this theme. Conversely, both groups' results were found to be normally distributed with regard to the data for the PU theme, although their results for the BI theme violated normality. In addition, the following boxplots were generated to check for any significant outliers to be excluded when running an independent samples t-test (see Figure 5.1).

Figure 5.1

Boxplot of Questionnaire Scores for the Intensive and Spaced Groups with Regard to the Three Themes, PEOU, PU and BI



Therefore, both parametric and non-parametric tests were run to analyse the quantitative part of the questionnaire (survey), based on the results of the assumption for each theme. The themes in the present study were based on Davis's (1989) three main components of the TAM framework, demonstrated by Dizon (2016) (see subsection 2.4.2.4). In the following subsection, the quantitative and qualitative questionnaire data will be analysed under each theme.

5.2.1 First Theme: Perceived Ease of Use (PEOU)

First, descriptive statistics were generated from the quantitative questionnaire data to obtain an overview of the students' perceptions of Quizlet's ease of use. In addition, the presentation of descriptive statistics helped to assess the data distribution, as well as evaluating the students' views of Quizlet's ease of use. Table 5.2 presents the mean scores for Quizlet's ease of use, as rated by the students, ranging from 1.00-5.00. The table also highlights other descriptive statistics, including minimum, maximum, mode, and standard deviation (*SD*), in order to give a clearer picture of the data.

Table 5.2

Descriptive Statistics for the PEOU Theme in the Experimental Groups

Group (N)	Range	Min-Max	Mode	Mean	*SD*
Intensive (47)	1.50	3.00, 4.50	4.38[a]	4.07	0.33
Spaced (40)	1.50	3.13, 4.63	3.75	3.80	0.37

[a] Multiple modes exist; the lowest value is shown

The Intensive Group's results for PEOU generated the higher mean score of 4.07. The Spaced Group also recorded a relatively high mean score of 3.80 for this theme. Therefore, it would appear that both these groups perceived Quizlet to be relatively easy and straightforward to use. However, there was a slight difference in the mean scores of the Intensive and Spaced Groups. Therefore, a further check for any differences between the Experimental Groups was performed, using inferential statistics. An independent samples t-test and the Mann-Whitney test were consequently run to reveal any differences between the groups. These two tests were applied, because the assumption of homogeneity was fulfilled, but the assumption of normality was partially violated (see Table 5.1). Moreover, a bias corrected (BCa) confidence interval was selected for analysing the independent samples t-test, because of the partial violation of normal distribution (Field, 2018). Thus, the results for the PEOU theme, generated using the independent samples t-test, indicated a significant difference between the Intensive and Spaced Groups, $t(79) = 3.67$, $p < .001$, and this difference, 0.28, BCa 95% CI [0.12, 0.43], represented a large effect size of $d = 0.82$. To verify this result, the Mann-Whitney U-Test was selected and also indicated a significant difference between the Intensive ($Mdn = 4.13$) and Spaced Groups ($Mdn = 3.75$): $U = 526.00$, $z = -3.54$, $p < .001$, $r = -0.38$, $\eta^2 = .15$. The effect size remained high on application of a non-parametric test (Mann-Whitney test). Therefore, the results of both the parametric and non-parametric tests showed that there was a significant difference between the Intensive and

Spaced Groups, the former reporting it to be easier to use, although both groups found Quizlet an easy website to use for learning vocabulary.

Consequently, the quantitative data quite clearly showed that the students found Quizlet very easy to use. The next step was to look at the qualitative data from 86 findings, collected from the Intensive (N= 46) and Spaced (N= 40) Groups. In the students' responses to open-ended questions in the second part of the Quizlet questionnaire, the results for the PEOU theme showed that most of the students in both groups (Intensive and Space) found Quizlet easy to use. For example, a student from the Intensive Group, Class 2, at No. 1 in the classroom list (Student IG2-1) stated:

I found Quizlet very easy and there were no difficulties.

Another student from the same class, Student IG2-9 declared:

Personally, I don't encounter any difficulties with the programme.

Meanwhile, a student from the Spaced Group, Class 3, at No. 13 in the classroom list (Student SG3-12) claimed:

It was beautiful, creative and easy to use. We enjoyed it.

However, it is interesting to note that not all the students found the programme easy to use. For example, two students from the Intensive Group (Student IG2-3 and Student IG2-6), representing 2.4% of the total findings from both groups, indicated that Quizlet was moderately easy to use. In contrast, there were two students from the Spaced Group (Student SG3-1 and Student SG4-19), also representing 2.4% of the total findings, who found Quizlet difficult to use. In addition, five students from both groups (representing 5.9%) mentioned some difficulties that they faced on starting to use Quizlet. For example, a student from the Spaced Group (Student SG3-2) stated:

Using Quizlet was a little bit difficult.

However, Student IG2-17 from the Intensive Group admitted:

It was easy, but I had a simple problem with opening an account.

and Student IG1-4 from the Intensive Group stated:

It was easy, except for some simple issues, which were managed and overcome following an explanation by the teacher.

It is therefore clear that most of the students found the programme easy to use. Despite the presence of students who experienced some difficulties with the programme, these represented a very small percentage. Therefore, both the quantitative and qualitative data support Quizlet as a vocabulary-learning tool that is easy to use from a student's perspective.

5.2.2 Second Theme: Perceived Usefulness (PU)

The second theme relates to how useful the students found Quizlet to be for their vocabulary learning. As illustrated in Table 3.2, there were nine statements in the questionnaire that assessed the usefulness of the software programme, based on the students' opinions. As with the first theme (PEOU), the statements on PU were rated using a five-point Likert scale, whereby the students indicated their views, ranging from 'Strongly agree=5' to 'Strongly disagree=1'. The descriptive statistics are presented in Table 5.3, summarising the students' views on the usefulness of Quizlet for vocabulary learning.

Table 5.3
Descriptive Statistics for the PU Theme in the Experimental Groups

Group (N)	Range	Min-Max	Mode	Mean	*SD*
Intensive (47)	0.78	3.56, 4.33	3.89	3.94	0.19
Spaced (40)	1.67	2.89, 4.56	3.67[a]	3.70	0.32

a. Multiple modes exist; the smallest value is shown

The descriptive statistics presented in Table 5.3 show high mean scores for the Intensive and Spaced Groups, of 3.94 and 3.70, respectively. This indicates that both groups considered Quizlet to be beneficial and enjoyable for learning new words. The Intensive Group also

attained a higher overall score than the Spaced Group, although the difference was quite small. In order to identify any differences between the two groups, the independent samples t-test was selected as the ideal test. However, this theme violated the assumption of homogeneity of variance ($p < .05$) (see Table 5.1). Such a violation can be corrected using the Welch-Satterthwaite method, which is presented in SPSS statistics as "Equal variances not assumed" (Statistics.laerd.com, 2019). Thus an independent samples t-test, using the Welch-Satterthwaite method was performed, which disclosed a significant difference between the Intensive and Spaced Groups: $t(61.13) = 4.07$, $p < .001$, with a difference of 0.24, 95% CI [0.12, 0.35], representing a large effect size of $d=1.26$. Although there was a significant difference between the views of the Intensive and Spaced Groups with regard to the usefulness of Quizlet, both groups valued Quizlet and found it beneficial and enjoyable for learning vocabulary.

As above, it is helpful to review the qualitative data, which showed that the students supported the use of Quizlet as a beneficial and enjoyable means of learning vocabulary. Almost all the students agreed that Quizlet was useful, with only one out of 85 students stating that they found it boring; this student, who was from the Spaced Group (Student SG3-1), stated:

Quizlet was difficult to use and boring.

However, all the other students mentioned that the use of multimedia meant that sound and images had a strong positive effect on their ability to memorise vocabulary quickly. They emphasised that presenting images alongside vocabulary helped them fully understand the meanings of words. They also found it useful to listen to the words being spoken, as a means of learning correct pronunciation. For example, Student SG3-3 stated:

I liked it that the images were displayed with vocabulary, and listening to the pronunciation of words. This was beneficial for understanding the meaning of the vocabulary and recalling it easily and quickly.

Meanwhile, Student SG4-2 declared:

Using images and listening to words helped me to recall the words quickly.

In addition, the students indicated the usefulness of receiving direct feedback from Quizlet, whenever they made any spelling mistakes or gave wrong answers. This allowed them to recognise their mistakes. The students also thought that it had helped them memorise words. To illustrate this, Student SG3-17 stated:

> *The thing I liked most about Quizlet was correcting mistakes, because that helped me to correct my own mistakes and not to repeat [them]. Also, repeating and hearing the words helped me to remember them, especially as I am an auditory [learner].*

Student IG2-8:

> *The programme gives the student an opportunity to correct mistakes, because this will help the student to remember and understand.*

The students also mentioned some Quizlet activities that they especially enjoyed and which they found particularly useful for learning vocabulary, such as Write, Spell, Test, Flash Cards and Match. However, the majority agreed that the photographic images used in Flash Cards were the most beneficial tools for learning new words. The students also indicated the Match game as an enjoyable activity. This required dragging the word to the correct photo, whereupon all the students' results appeared, revealing those who had accomplished this task in the least amount of time. Some of the students pointed to the importance of providing L1 translation, especially the Level 1 students. This was pointed out by Student SG4-3:

> *The most useful activities were in Flash Cards, because the images and words helped me to memorise the words very quickly, and also the Match game, because it created competition between students.*

and Student SG4-17:

> *I liked learning vocabulary using sound, images, L1 translation, and challenging games like Match, which provided an atmosphere of enthusiasm and a spirit of challenge.*

To sum up, the overwhelming response from the students in both groups was that they valued Quizlet, and this was evident from both the quantitative and qualitative data generated. The students considered Quizlet to be useful, enjoyable and beneficial for learning and recalling vocabulary. However, the quantitative data indicated that the Intensive Group

demonstrated a greater tendency to value Quizlet as a useful tool for learning vocabulary. In addition, all the findings from the qualitative data from both groups were positive, except for one participant from the Spaced Group, who found Quizlet boring.

5.2.3 Third Theme: Students' Behavioural Intention (BI) for Future Use

There were eight items (statements), for the third theme, BI, categorised as possible interest for use again in the future. Here, a five-point Likert scale was similarly applied to measure the students' willingness and intention to use Quizlet in future. The first step was to report the descriptive statistics (see Table 5.4) and then compare the results for both groups (Intensive and Spaced), using the Mann-Whitney test to check for any differences between their responses.

Table 5.4
Descriptive Statistics for the BI Theme in the Experimental Groups

Group (N)	Range	Min-Max	Mode	Mean	*SD*
Intensive (47)	2.00	2.50, 4.50	3.50	3.32	0.34
Spaced (40)	2.00	2.00, 4.00	3.13	3.21	0.34

The descriptive statistics presented in Table 5.4 indicate convergence between the results of the Intensive and Spaced Groups. The mean scores for both groups were very close, with the same range (2.00) and standard deviation (0.34), and a mean rating of 3.32 and 3.21 for the Intensive and Spaced Groups, respectively. Therefore, both groups appeared to agree that they would use Quizlet for their future vocabulary-learning.

Because the assumptions of normality and homogeneity were violated for the BI theme (see section 5.2 and Table 5.1), the Mann-Whitney test was run to evaluate the differences between responses to the five-point Likert scale items. The test revealed no significant

difference between the Intensive (*Mdn*= 3.38) and Spaced Groups: (*Mdn*= 3.19), *U*= 719.00, *z*= -1.90, *p*= .057, *r*= -0.20, η^2= .04.

The qualitative data were then analysed to identify the students' views on using Quizlet in future. The results for the BI theme showed that all the students, except for one, planned to use Quizlet in future. Below are several examples of students' responses to the question of whether they would want their College to use Quizlet for vocabulary learning:

Student IG1-8:

> *Strongly yes, because it fixes the words into the mind.*

Student IG1-12:

> *Yes, for sure, because it is beneficial and important for students. It is an easy programme for learning the English language.*

Student SG3-10:

> *Yes, Quizlet is the best out of all the programmes I have used to learn English so far. It will help students to learn new words.*

This willingness amongst almost all the students in both groups to continue using Quizlet was a powerful indication of their satisfaction. Both the qualitative and quantitative findings from the questionnaire supported the overall perception of Quizlet as a straightforward and effortless website. The findings also showed that most of the students considered the programme uncomplicated. In addition, the majority of the students found Quizlet valuable and helpful for learning vocabulary. Therefore, the students had a positive opinion of Quizlet's ease of use and usefulness. Consequently, the students in both groups (Intensive and Spaced) demonstrated the intention to use Quizlet in future. The following section will now reveal the views of the teachers, with regard to using Quizlet for learning vocabulary.

5.3 Interviews: The Teachers' Perceptions

The second set of findings in this chapter consists of the qualitatively analysed interviews with core teachers in the Experimental Groups (Intensive and Spaced). These interviewees comprised a small group of four teachers. As such, the researcher acknowledges that no attempt at generalisation can be made. However, these interviews were conducted for two main reasons. First, the current research is practice-based and so it was important to gather the views of those teaching the students and supporting their learning, as well as the views of the students themselves. In addition, the teachers' views were important for obtaining more information about Quizlet when used on a day-to-day basis, as well as to see if they differed from those of the students in the quantitative data. The teachers consequently responded to 10 questions, grouped under the three main themes (PEOU, PU and BI), as shown in the following subsections.

5.3.1 First Theme: Perceived Ease of Use (PEOU)

Almost all the teachers agreed that Quizlet was not complicated for students to use and that they found it convenient, as the students merely needed to follow the researcher's instructions for registration, logging in, and using the programme. Intensive Group Teacher Two (IGT-2) observed that:

> *Logging in was easy and smooth for most of the students. Implementation was perfect during the class… amazing programme that kept the students fully engaged with its different features.*

When this teacher was asked about the clarity and simplicity of the L1 translation and photographic images provided in the programme, she affirmed:

> *Yes, clear enough and meaningful.*

Correspondingly, Spaced Group Teacher One (SGT-1) declared that the programme was as easy for her to use, as a teacher, as it was for the students:

> *For me as a teacher, the programme was easy to use and not complicated. For the students, they didn't face any difficulties in accessing the programme or in logging in or*

out. The activities were also designed to be easy and useful… the students didn't complain of any obstacles.

However, Spaced Group Teacher Two (SGT-2) emphasised that teachers and students need training in using the programme, which would make it easy to use:

Ok, like any of the programmes we use with our students, when it comes to technology, it takes time for them to get used to it, of course, and so that requires training. It would have to be, I would say, maybe two or even three sessions to make sure that they understand all aspects… Also, teachers would have to receive training, but this is totally doable… I think it would be easy to use, once everybody understood how to use it.

Therefore, the qualitative data from the teachers' interviews revealed positive views of Quizlet, based on its ease of use. The results corresponded to those gathered through the students' questionnaire. Both the teachers and students agreed that the programme was easy, clear and uncomplicated to use.

5.3.2 Second Theme: Perceived Usefulness (PU)

Several questions in the interview (see subsection 3.4.4.1 and Appendix 9) were raised to determine whether the teachers felt that Quizlet was beneficial. These questions centred upon its usefulness, the students' progress and motivation, and any aspects that were especially or least useful in Quizlet. Here, all the interviewees agreed that Quizlet was enjoyable and beneficial for students to use. For example, SGT-1 explained:

Yes, it is useful, because the student will be able to see a picture and match it with its meaning, and listen to the pronunciation of words… this will actually increase his knowledge and of course, help him to expand his vocabulary with the right pronunciation in an interesting way… without a doubt, the students showed obvious progress in their vocabulary knowledge, because they learned through vision, listening and writing the words in an interesting way.

SGT-2 also mentioned that the students enjoyed using Quizlet more than Moodle (see subsection 1.2.2), an online platform that had already been used at the College. She thought that the students were more fully engaged by Quizlet, because it was interesting and practical:

SGT-2: Personally, I did not use it, but when I saw the students using it, they were really engaged.

...I have to say that Quizlet looks like a really great way for students, not only to learn vocab, but also other aspects of English during lab time. This would be far better than Moodle, for example: something practical, something they can enjoy.

Researcher: The Moodle that is used at the College?

SGT-2: Yeah, exactly.

In addition, all the interviewees shared the same view that Quizlet motivated students to learn vocabulary. Intensive Group Teacher One (IGT-1) claimed that the students

...improved their vocabulary. They liked it. They were motivated to use it…

while IGT-2 mentioned that she had noticed some progress in her students, commenting:

Yes, because my students remembered most of the words that they had learned using Quizlet.

However, the teachers held different views on the aspects of the programme that they found most or least useful. For example, IGT-1 considered the task of creating a vocabulary set to be the most useful, but found providing students with L1 translation to be the least useful, expressing:

The part I found least useful was the use of bilingual translation. I prefer using images, context and synonyms.

This opinion differed from those of all the other teachers, who considered L1 translation to be useful for the students, with SGT-2 stating:

I think for my personal… from my own learning and from seeing how students learn, I think L1 translation's really important. I really do, and I use it with my students too.

In fact, SGT-2 found the Write activities to be the most useful, because these involved the highest rate of word production by the learners. However, she believed the Match game to be the least useful, declaring:

Least useful… well, even though they enjoy it, maybe the matching is the least useful, in the sense that it's not a production skill, so it might be the opening of the door for them.

However, IGT-2 held the completely opposite view to SGT-2, as regards the most and least useful Quizlet activities, remarking:

The Match Game was the most useful based on the feedback I got from my students, because it is fun and more visual. Write is the least useful, because it is very challenging for the students.

Moreover, SGT-1 cited various useful aspects of Quizlet, whereby teachers and students could select whatever they thought would meet their needs:

Well, from my own point of view I found all aspects useful… the pronunciation of English words was very useful, because it got the students to achieve the exact pronunciation of each word; also, the writing part helped students to practise writing words.

To sum up, all of the teachers considered Quizlet to be a very useful tool for vocabulary learning, although they differed on which specific aspect was most useful. These qualitative data supported the students' views, as gathered by the questionnaire. Therefore, almost all the participants (teachers and students) strongly agreed that Quizlet was beneficial and enjoyable.

5.3.3 Third Theme: Teachers' Behavioural Intention (BI) for Future Use

The teachers were asked if they planned to use Quizlet in the future with their students. Additionally, they were asked if they would recommend the programme to other teachers. The results showed that all of the teachers wished to use Quizlet and were willing to recommend it to their colleagues for use in the classroom. For example, when SGT-3 was asked if she would recommend the programme to others, she declared:

Without a doubt, yes, for the reasons already mentioned and because Quizlet is a programme that can provide students with different methods of learning in an exciting way.

IGT-2 had encouraged the students to install Quizlet on their mobile phones, in order to facilitate their vocabulary learning. She emphasised that students

should install it on their mobiles and keep learning new words all day.

Similarly, IGT-1 stated that

students should be encouraged to use Quizlet outside the classroom. They should be instructed to install the application on their smart devices, as this will help them improve their vocabulary knowledge.

In conclusion, all the teachers and almost all the students were willing to use Quizlet for vocabulary teaching and learning.

5.4 Summary of the Questionnaire and Interview Findings

To condense, the quantitative and qualitative findings for the students' completed Quizlet questionnaires were utilised to answer the first and third research questions, while the qualitative findings from the interviews conducted with the teachers were used to answer the third research question. Overall, the findings from the Quizlet questionnaire and interviews indicated that the participants (the teachers and students from both the Intensive and Spaced Groups) held positive views of Quizlet in terms of its PEOU, PU and their BI to use it.

Examining the quantitative results of the Quizlet questionnaire by running the independent samples t-test and Mann-Whitney test, the Intensive Group was more positive than the Spaced Group about Quizlet's usefulness and ease of use. However, both groups were equal in their intention to use Quizlet. The students' qualitative data (gathered using open-ended questionnaire items) and the teachers' interview data supported the quantitative data collected using the questionnaire, thereby indicating that almost all the participants, comprising students and teachers in both groups, had positive perceptions of using Quizlet, with regard to its PEOU, PU and their BI to use it in future.

In the next chapter, the experimental learning gains will be discussed to investigate the impact of lag effects (longer versus shorter intervals between practice sessions), using two different time sequences (intensive versus spaced) to support vocabulary learning and retention in low ability learners, and to explain students' and teachers' perceptions of the use of Quizlet. The Discussion Chapter is intended to answer the research questions in light of spaced effect theories applied to L2 vocabulary learning and Davis's (1989) TAM framework.

CHAPTER SIX: DISCUSSION

6.1 Introduction

The central focus of the present study was to improve vocabulary learning and retention amongst students with a low level of English proficiency, within a typical classroom setting. A quasi-experimental design was consequently implemented to examine the effectiveness of time distribution (intensive versus spaced) on vocabulary learning, using a Computer-assisted Language Learning (CALL) programme (Quizlet). Initially, the sample consisted of 143 participants (139 students; 4 teachers). However, only 96 of the students attended all the sessions and took all the tests. This is because the current study was conducted in an actual college and not in a laboratory, and so may be closer to the reality faced by teachers in their day-to-day practice, during which student absences are common.

In addition, this study investigated students' perceptions of the use of Quizlet to help them learn English. It also enquired into teachers' perceptions of the learning tool, in relation to the most important elements of their course design (Graves, 2000). Therefore, this current chapter discusses the main findings presented in the two previous chapters (Chapters Four and Five), with reference to the theories and existing literature that are presented in Chapter Two. The discussion of these key findings is organised and summarised according to the three primary research questions (RQ) addressed in this study:

RQ1. Is Quizlet an effective programme for promoting vocabulary learning amongst low ability learners in the classroom?

RQ2. Does the time distribution of the practice sessions (intensive versus spaced) moderate the benefits of using Quizlet to promote vocabulary learning and retention amongst low ability learners?

RQ3. What are the perceptions of teachers and students regarding the use and implementation of Quizlet?

6.2 RQ1. Is Quizlet an Effective Programme for Promoting Vocabulary Learning amongst Low Ability Learners in the Classroom?

In order to examine the effectiveness of using Quizlet to promote learners' vocabulary knowledge, three performance tests were administered, and their results compared between the Experimental Groups (Intensive and Spaced) and the Control Group. At first, all three groups (Intensive, Spaced and Control) took baseline tests (VLTs and WM tests) to check for any significant differences between them. The Kruskal-Wallis test for the 2,000, 3,000 and 5,000 VLTs subsequently revealed no statistical differences between the groups, with small effect sizes. For example, the groups demonstrated a similarly low level of vocabulary knowledge, falling below the 2,000-word frequency (see subsection 4.4.1). Moreover, one-way ANOVA for the WM tests revealed no statistically significant differences between these groups in their ability to retain vocabulary (see subsection 4.4.2) before the intervention. Therefore, all three groups possessed the same level of vocabulary knowledge and ability to recall information.

'Effectiveness' implies that using the CALL tool (Quizlet) had a positive effect on the acquisition of new vocabulary. The results of the descriptive statistics and two-way mixed ANOVA showed that the technology appeared to be helpful for vocabulary learning (see section 4.5). This is because both the Experimental Groups (Intensive and Spaced) using Quizlet made highly significant progress, which is in line with Experimental Hypothesis 1.1 (see section 3.2). Furthermore, they greatly outperformed the Control Group in the immediate post- and delayed post-test, whilst the Control group did not progress. In fact, the Control Group's results showed no differences at any of the three time points, whether at short-term (4-day RI) or long-term intervals (28-day RI).

It should be noted that the Control Group were a test-only group and did not receive any treatment sessions during the intervention period. Instead, they continued to receive their regular teaching, with no instruction involving the target vocabulary. As mentioned above, there were no changes in the Control Group's results, while the Experimental Groups made significant progress. This result was expected, but it is a key outcome, because it means that there was limited influence and interference from other potential factors – such as the test effect, the course books, or the teachers' instruction during the intervention, all of which could have led to improvement by means other than using Quizlet. It implies that the

improvement in the results obtained by the Experimental Groups is most likely attributable to the treatment that they received during the intervention, a finding which could not be confirmed without the presence of a Control Group. Hence, the Control Group helped the researcher to ascertain that it was specifically the intervention that had a positive effect on students' results in the Experimental Groups.

Aside from the above, the findings of this study reflect the positive outcome of using Quizlet in both the short and long term. These findings are in line with those reported by Özer and Koçoğlu (2017), who compared two experimental groups (a Quizlet group and Notebook group) with a control group. Özer and Koçoğlu's findings from a study on vocational High School students demonstrated that the Quizlet group significantly outperformed the control group and also performed slightly better than the Notebook group in post- and delayed post-tests. As far as the current researcher is aware, Özer and Koçoğlu's study is the only empirical evaluation of the effectiveness of Quizlet, which looks at the long-term gains in a quasi-experimental approach at three time points. There are in fact very few empirical studies (for example, Korlu & Mede, 2018; Sanosi, 2018) that have evaluated the effectiveness of Quizlet, and these have either deployed just two performance tests (pre- and post-tests) (Sanosi, 2018), or administered two performance tests with a survey (Korlu & Mede, 2018; Zambrano Acosta, 2018). Meanwhile, other studies have merely addressed perceptions, rather than learning effectiveness (for example, Anjaniputra & Salsabila, 2018; Chien, 2015; Jackson III, 2015). Therefore, this present study makes an original contribution to knowledge by attempting to fill the gaps in the existing literature by evaluating students' long-term retention via a delayed post-test. It also endeavours to bridge the gap between theory and practice, as this study is practice-based, adopting an explicit method of vocabulary-teaching via Quizlet, under two different spaced practice schedules to promote vocabulary-learning in a real world setting.

6.2.1 The Effect of a CALL Tool in Practice

A quasi-experimental design was chosen to ascertain whether Quizlet was an effective learning tool, particularly in the context of classroom-based learning. This was very important for demonstrating that the tool could be deployed in practice. It indicates that the present study investigated an effective tool, which helped students to learn and retain vocabulary in

the long term, and not in a laboratory setting, but rather in their typical classroom. The Experimental Groups' outperformance of the Control Group revealed Quizlet to be highly effective. Therefore, the results of this study could help the College to find a different way of supporting its students in their English language learning.

Aside from the above, the present study findings show that Quizlet, which is freely available in schools and other educational settings, can have a sustained impact on vocabulary knowledge amongst learners who have not been able to progress beyond the pre-elementary level of language learning. The remarkable progress of the Experimental Groups (Intensive and Spaced) reflects the benefits of using Quizlet for vocabulary teaching and learning amongst learners with a low level of English proficiency. Therefore, Quizlet may be considered effective for teaching vocabulary to learners with low ability in this area. It is a finding that supports Dizon (2015), who demonstrated that Quizlet was an effective tool for vocabulary learning, although Dizon did not conduct a delayed post-test and the sample size was very small (see subsection 2.4.2.4). Therefore, the current study makes an important contribution to knowledge, as it provides statistically significant evidence, supported by large effect sizes, which demonstrate that these groups made substantial progress in the immediate term. Moreover, the progress was sustained, according to the results of delayed post-tests in a typical classroom setting (see section 4.5).

Beyond the quantitative data, however, the qualitative data also pointed to progress. For example, one teacher (SGT-2) intended to be using Quizlet instead of Moodle (see subsection 5.3.1); it was applied once a week in a session that is currently used for Moodle (an online platform provided by the College; see subsection 1.2.2). All the teachers participating in this study valued Quizlet and found it useful for students' vocabulary learning. In addition, the students in the Experimental Groups held positive views of Quizlet, according with the findings of previous studies (for example, Korlu & Mede, 2018; Zambrano Acosta, 2018), where it was found that the tool benefited students' vocabulary learning (this point will be covered in more detail in section 6.4). The quantitative and qualitative data make a key contribution to the research, showing Quizlet to be beneficial, with statistically significant progress being possible amongst this specific cohort of students in a typical classroom setting at the College.

6.2.2 Selecting the Study Sample

Although there is a great deal of research that looks at progress in vocabulary learning, the relevant literature is largely focused on more able learners (for example, Anjaniputra & Salsabila, 2018; Dizon, 2016; Korlu & Mede, 2018; Lander, 2016). However, some students do not fit into the category of more able language learners, but rather have not been able to progress beyond the pre-elementary level of language learning, and are often ignored by the research community. In fact, very few recent studies have been conducted using lower ability learners to evaluate the effect of Quizlet on vocabulary learning at university level (for example, Barr, 2016; Sanosi; 2018). This is a serious gap in the literature and, as argued in this thesis (see section 2.4.2.3), a failure to give attention to this topic has meant that no guidance is provided to teachers on effective teaching methods and strategies. However, teachers need to know how to meet the needs of this particular cohort of students, so that they can make progress in their vocabulary knowledge. Consequently, teachers and researchers are responsible for supporting any of the learners involved, including those group members who fail to progress, and not just those who find it easy to learn English. It is easier to teach students who understand what they are taught, but it is also important to find appropriate methods to help learners who find it difficult to understand everything immediately. In fact, all the students in this study were at the pre-elementary stage. As a result, this study aimed to contribute to the existing knowledge by identifying the impact of Quizlet in practice, particularly in classroom-based learning and with a cohort of students who did not find it easy to learn English.

It is important to learn English, as it is a key skill that is often required by High School students worldwide. In this regard, Oman is not unique. The expectation and even requirement in many Secondary or High Schools around the world is that their students will be proficient in English by the time they graduate. Globally, every sector, from tourism to the rest of the business world, displays a preference for employees with English language skills, meaning that individuals with such knowledge have a significant advantage in the job market over those who do not (for a detailed review of the literature on this point, see section 1.2.1). However, this is not easy for every student to achieve.

That said, although English is considered to be a global language, which is spoken and sought after around the world, it does not necessarily make it easy to learn. In fact, some learners find it very difficult, especially native Arabic speakers (see subsection 1.2.3). For some students, traditional teaching methods are effective and helpful. However, this is not the case with every learner. Hence the question remains, what should be done with students who cannot learn by traditional means? The reasons why this particular group of learners failed to make progress before beginning their College studies could be due to factors within and beyond the school. The present study acknowledges that a variety of reasons may underpin this problem; however, detailed discussion of these reasons is beyond its scope. Instead, it seeks to find ways to support learners who are still at the pre-elementary stage, despite many hours of English teaching at school. As a result, the present study contributes to existing knowledge by showing that this particular group can indeed make progress using Quizlet, a freely available software application. The fact that it is a free software application is an additional advantage of Quizlet, making it more accessible to schools around the world, in contrast to paid software applications. As such, cost will not be a barrier for schools, colleges or universities in offering this support to students who have not progressed beyond the pre-elementary level of language learning.

6.2.3 The Effect of the CALL Tool on Students' Motivation and Engagement

This study showed that by using the Quizlet computer software package, the students in the Experimental Groups were motivated, and considered Quizlet to be an interesting learning tool (more explanation and details of the students' views of using Quizlet may be found in section 6.4). The positive impact of Quizlet on students' motivation and engagement was also reflected in the Experimental Groups' engagement with the post-tests. At both the immediate post- and delayed post-tests (for the 4-day RI and 28-day RI sub-groups), a significant reduction in the number of missing responses (i.e. the number of times that the students omitted to answer) in the Experimental Groups' results was observed. In contrast, the results from the Control Groups showed no change in the number of missing responses – these remained the same over the three time points (see subsection 4.5.1). In addition, there was a decrease in the number of spelling errors made by the Experimental Groups, with no major changes between the immediate post-test and delayed post-test (see subsection 4.5.1). This is a rather interesting finding, because it implies that following the intervention, the students had

more confidence to attempt a greater number of questions, and their spelling had improved. Therefore, the researcher speculates that technology is an exciting and innovative way to support language learning. The results derived from the missing responses also showed that technology could help the particular type of students included in the sample to stay engaged for longer. The results were in line with those reported by Tam et al. (2010), who considered CALL tools to be capable of creating an ideal environment for learners with low English language proficiency. Similarly, Krish et al. (2011) take the view that online tools can help learners build their confidence and increase their level of motivation.

However, the present researcher also acknowledges that the precise reasons behind enhanced student engagement when learning vocabulary with Quizlet are not known. One explanation could be that the students were interested in working with computers, since they had mostly grown up highly computer-literate. Therefore, using a computer was a typical part of life for them. It could equally be because the typical sessions involving work on a computer corresponded to this particular group's strengths in digital literacy skills, which promoted their learning. Nevertheless, it is beyond the scope of this study to ascertain the specific reasons. Moreover, this study does not compare Quizlet with any other online learning tools. As such, it cannot claim that Quizlet is one of the most effective software tools for learning vocabulary, but simply affirms that Quizlet is an effective vocabulary learning tool for students with low English language proficiency, who had not previously been able to make progress in a traditional face-to-face classroom teaching context. This was demonstrated in this study, whereby the Experimental Groups made some sustained progress. It was clear from the Experimental Groups' results on using Quizlet that they were more motivated than the Control Group, who were taught using traditional means.

6.2.4 The Effect of Explicit Vocabulary Learning

The progress made by both Experimental Groups perhaps appears to confirm the importance of explicit vocabulary teaching. In other words, the study supports that vocabulary learning and retention may be achieved through explicit teaching in an intervention using Quizlet. In this study, improvement was associated with enhanced recognition and vocabulary recall, linked with the theory of explicit vocabulary learning (see subsection 2.3.1.1). This is in line with previous studies (for example, Hunt & Feng, 2016; Nation & Meara, 2010), which

support explicit vocabulary teaching. For instance, Hunt and Feng (2016) advocate direct teaching, using reading strategies and multimedia to develop vocabulary knowledge. Meanwhile, Lin (2010) highlights the benefits of multimedia vocabulary software programmes to enhance learners' vocabulary knowledge. The present study demonstrates that the Quizlet programme, which includes multimedia aids, such as sound and images, can help learners recognise meanings and recall the forms of target words.

Nation and Meara (2010) encourage the use of explicit vocabulary teaching with learners at the earliest stages of their language learning. However, Kwiatkowska (2007) emphasises that when teaching vocabulary, it is important to implement an interesting and effective tool that is appropriate for the learners' level. The present study showed a high degree of progress made by the Intensive and Spaced Groups, who had low English language ability. Therefore, the beneficial role of explicit learning via the use of Quizlet was demonstrated, thereby contributing to the existing literature that supports explicit vocabulary instruction. It should be added that Quizlet proved to be effective in promoting vocabulary learning, because it supported direct instructional methods of teaching vocabulary in practice, with the specific cohort of learners selected for the study sample.

6.2.5 Summary of Discussion for RQ1

In response to the first research question (RQ1), the findings of the present study showed that the use of Quizlet in the classroom was effective for facilitating vocabulary learning amongst learners with low English language ability. This study focused on what happens in a typical classroom, given the lack of available literature in this field; there is no actual application of Quizlet to teaching and learning practice in the existing research. Thus, the present study showed that it was possible to implement the application in practice.

Furthermore, the study groups' scores indicated that the Experimental Groups outperformed the Control Group, thereby demonstrating the effectiveness of using Quizlet in this particular College, with students who were not progressing in their English language learning. In addition, the participants in the Experimental Groups showed their motivation and engagement in vocabulary learning, both in their reflection on the use of Quizlet and in the reduction in their missing responses over the three test time points. Finally, the significant progress made in the Experimental Groups' scores support explicit vocabulary learning as an

effective strategy for learners with low ability in this area. Even though they had been learning English for a long time, going back to Year 1 in school (see subsection 1.2.2), they still had low English language ability and continued to be unable to progress beyond the pre-elementary level of language learning.

6.3 RQ2. Does the Time Distribution of the Practice Sessions (Intensive versus Spaced) Moderate the Benefits of Using Quizlet to Promote Vocabulary Learning and Retention in Low Ability Learners?

As mentioned above, this study focused on learners with low English language ability in actual classroom practice, in an attempt to assist them in their vocabulary learning and retention. Further, the impact of the lag effect on vocabulary learning has been compared, i.e. longer versus shorter spacing between practice sessions, in order to investigate the long-term retention of target vocabulary. To discover this effect of short spacing (1-day ISI) and long spacing (7-day ISI) for this particular group of students, two-way mixed ANOVA and *post hoc* pairwise comparisons, using Bonferroni corrections, were also utilised (see section 4.5). A comparison was made between the results of the two main sub-tests (Recognition and Recall) at two different retrieval intervals (RI), i.e. in delayed 4-day RI and 28-day post-tests. Each main sub-test was further divided into Passive and Active sub-tests (i.e. Passive/Active Recognition; Passive/Active Recall).

The remarkable finding for the Recognition and Recall sub-tests was that both the Experimental Groups (Intensive and Spaced) had improved significantly by the immediate post-test, sustaining this improvement in both delayed post-tests (4-day RI; 28-day RI). This finding is confirmed by very large effect sizes, as well as confidence intervals that do not cross zero, which means that the findings are reliable. It also provides good confidence in the data and findings. Nevertheless, it should be noted that there was a significant decline in the Recall sub-test results for the Intensive Group in both the delayed 4-day RI and delayed 28-day RI post-tests, as well as for the Spaced Group in the delayed 28-day RI post-test. This could be due to the learners being more capable of retaining receptive than productive knowledge (Schmitt, 2010). However, these findings may not reflect a reliable effect, because of very small effect sizes and confidence intervals crossing zero, which indicate that the scores for the delayed post-test had only decreased by a small amount and remained higher

than at the baseline. These findings indicate that both the short spacing (1-day ISI) and long spacing (7-day ISI) between practice sessions for the Intensive and Spaced Groups, respectively, were sufficient for recognising meaning and producing the forms of target words (see subsection 6.3.3 for more discussion on this finding and its effect on the vocabulary-learning process for students, teachers and schools).

Turning now to RQ2, the Intensive (1-day RI) and Spaced (7-day RI) Groups, with their two different spacings, were found to progress equally between the pre- and immediate post-tests and to have maintained their progress in the delayed post-tests. This finding suggests that Experimental Hypothesis 2.1 (that gives the preference to the Intensive Group in the delayed 4-day RI post-test) and Experimental Hypothesis 2.2 (that supports the Spaced Group in the delayed 28-day RI post-test) should be rejected (see section 3.2). Therefore, both Experimental Groups appeared to have received great benefits from using Quizlet to enhance their vocabulary learning and retention; this is a particularly important finding for these participants, who had not previously made progress in their English language learning. Further, these benefits do not seem to have been influenced by the length of time between practice sessions.

6.3.1 Evaluating the Optimal Spacing

The finding that both the Intensive and Spaced Groups made equivalent gains was interesting, as it was unexpected to find no statistically significant differences between the Intensive and Spaced Groups, in terms of the acquisition (immediate post-test) and retention (delayed post-test) of target vocabulary. This is because the present study attempted to utilise the optimal spacing suggested by Rohrer and Pashler (2007), whereby the optimal ISI should be between 10% and 30% of the RI range. Therefore, the researcher extended the design of Serrano and Huang's (2018) study, using two delayed post-tests with the same optimal spacing for each Experimental Group (1-day ISI/4-day RI=25%; 7-day ISI/28-day RI=25%). In Serrano and Huang's study, it was concluded that shorter spacing supports immediate progress in vocabulary learning, whilst longer spacing sustains long-term retention (see subsection 2.4.1). However, these findings do not correspond with those of the present study, which showed no statistically significant differences between the Intensive and Spaced Groups in the short- and

long-term, delayed post-tests (immediate – 0-day RI; 4-day RI; 28-day RI). Further, both groups retained their learning gains at both the 4-day and 28-day RI delayed post-tests.

It was expected that the Spaced Group would do better in the 28-day RI, because there were longer periods of time between their practice sessions. This is because introducing a longer delay between practice sessions can make retrieval of previously learnt information more effortful, and effortful retrieval is thought to strengthen long-term memory (Toppino & Gerbier, 2014). Therefore, the Spaced Group were predicted to retain more knowledge by the post-test due to the longer delay (7 days) between their practice sessions. However, both groups in the current study performed equally well in the immediate post-test and maintained this performance in both the short- and long-term, delayed post-tests. This finding contradicts Rohrer and Pashler's prediction that the interval between the final practice session and the testing session needs to fall within the optimal time period, i.e. the Intensive Group should do better in the 4-day post-test and the Spaced Group should do better in the 28-day post-test. Contrary to this prediction and the findings of Serrano and Huang (2018), the results of this study suggest that both the shorter and longer spacing led to successful learning and retention of the target vocabulary.

It is also worth acknowledging that there are key methodological differences between the current study and Serrano and Huang's study.

First, the present study divided each Experimental Group into two sub-groups, each with a different delayed post-test; the first sub-group was administered the delayed post-test after four days, and the second took the delayed post-test after 28 days. Therefore, data was gathered from both Experimental Groups at both delayed post-tests (4-day RI; 28-day RI), but with different participants in each test. Conversely, in Serrano and Huang's study, all participants in each Experimental Group received the same delayed test, which matched the optimal spacing. This means that the Intensive Group completed the delayed post-test at 4-day RI, while the Spaced Group completed the delayed post-test at 28-day RI to match the optimal spacing of each group. Serrano and Huang's (2018) study, therefore, neglected to compare the two Experimental Groups in the same delayed post-test. However, this limitation was resolved in the present study, as the current researcher was able to assess both the Intensive and Spaced practice distribution in terms of the shorter and longer retention periods.

Secondly, like most of the previous literature on the lag effect, Serrano and Huang's (2018) study failed to include a control group. The deployment of a control group could have helped to ascertain whether the students' progress in the intensive group was due to the teaching method or test effect. In fact, the progress of the intensive group in Serrano and Huang's study could be attributable to the test effect, which is not known, because there were only four days between the immediate post-test and the delayed post-test for the intensive group. For example, in the current research, there was a significant change found for the Control Group in the delayed (4-day RI) post-test, regarding the results of the Active Recognition sub-test. This could reveal that the test effect was due to the short interval of just four days between time points 2 and 3. (However, the effect size was very small, with the confidence interval crossing zero (see subsection 4.5.4.1), suggesting that this finding is unreliable.) Nevertheless, the existence of a control group, together with the effect size, was important to confirm whether the progress observed within a shorter RI was because of the test effect. To the best of the current researcher's knowledge, only one very recent lag effect study has included a control group (Kasprowicz et al., 2019), but this focused on the impact for L2 grammar learning with young learners, which is beyond the scope of the current study. Therefore, the present study has attempted to fill this gap in the literature by deploying a control group.

These methodological differences between the present study and Serrano and Huang's study make it difficult to draw conclusions about the reasons behind the differences in results between the two studies.

Almost all the literature on the spaced effect concurs that learners need intervals between practice sessions. This is also supported by the pilot study in this current research. The findings from the pilot study showed that the spaced group outperformed the massed group (see subsection 3.4.5 for more details). In addition, the previous literature on massed versus spaced practice supports the importance of spacing the practice sessions (Kornell, 2009; Lotfolahi & Salehi, 2017). However, what has not yet been agreed on in previous lag-effect studies is how long this space should be. To be more precise, how much time should elapse between the practice sessions? In fact, the findings of these previous studies, which looked at different intervals (one day, three days and seven days) were very mixed (see subsection 2.4.1). Additionally, the length of these intervals proved to be of minor importance

for the learners in this present study, as both groups produced equivalent results. Consequently, these findings suggest that there might not be an optimal interval, so much so as simply a need to space practice sessions. This is also where differences between individual learners could be significant (Larsen-Freeman, 2014), because some learners may prefer an interval of one day, while others might prefer seven days. This factor would influence whether or not intervals of different lengths are effective and lead to the mixed findings between studies. Therefore, the research to date, including the present study, indicates that it is important to have space between practice sessions, but it is not possible to draw any conclusions about the optimal length of that spacing.

6.3.2 Lag Effect in Practice

The findings of this current study contradict the previous literature (for example, Bird, 2010; Rogers, 2015), where it was found that a longer spacing of seven days or more was much more beneficial for retention. It also contrasted with studies (for example, Suzuki, 2017; Suzuki & DeKeyser, 2017), where it was found that shorter spacing between practice sessions (1-day to 3.3-day ISI, respectively) could yield better recall (see subsection 2.4.1). However, all these above-mentioned studies were laboratory-based, and focused on L2 grammar learning, which differs from the focus of the present study. Suzuki (2017) attributes these mixed results to the influence of aspects relating to L2 skills. However, the present study aligns with Kasprowicz et al.'s (2019) study, where it was found that short intervals of 3.5 days and long intervals of 7 days were equally beneficial for young learners in terms of acquiring French morphosyntactic structures. Kasprowicz et al.'s study was quasi-experimental and implemented in an ecologically valid classroom context. Therefore, the mixed results could be the result of the setting. For example, results can differ according to whether the study context is a real-life setting (for example, the classroom) or a clinical study setting (such as a laboratory). Thus, Küpper-Tetzel et al. (2014) emphasise the importance of further research in real classroom settings, as this will expand the validity and evidence of the impact of memory on real-life learning environments. Therefore, the present study contributes to the existing knowledge by identifying the length of spacing required for the long-term retention of knowledge in classroom-based learning, specifically with a group of students who did not find it easy to learn English.

There have in fact been very few empirical studies (see Küpper-Tetzel et al., 2014; Rogers & Cheung, 2018; Serrano & Huang, 2018) conducted in a classroom setting, which have looked at the lag effect in relation to vocabulary acquisition. With regard to findings for the long term, the results of this current study are generally in line with those obtained by Küpper-Tetzel et al. (2014), namely that both short and long spacing of practice sessions produce the same learning outcomes. Conversely, the findings of the present study contradict those of Rogers and Cheung (2018), who support that shorter spacing between sessions leads to long-term retention (see subsection 2.4.1). However, these differences could be due to different settings and contexts, as the present study was conducted with adults in their first year at college, whilst the participants in previous studies were younger learners at Primary or Secondary school. The type of tasks and number of practice sessions also differed. Therefore, more research is needed to form a clearer picture of the impact of the lag effect in the same authentic learning settings.

As mentioned previously, both groups in the present study made the same progress. Consequently, it appears that it was important for both groups of this cohort of students to be given space between practice sessions, whether for one day or for seven days. It seems that this spacing facilitated the integration of the information into the short-term memory and then consolidated it into the long-term memory. This could be attributed to the spacing between the practice sessions, which helped support the working memory in transferring the information into the long-term memory, regardless of whether this spacing was one or seven days (Cepeda et al., 2006; Rohrer & Pashler, 2007). In addition, both groups had repeated opportunities to rehearse the new knowledge, which seemed crucial for the long-term memory. This will be discussed in more depth later in this chapter (see subsection 6.3.4.).

6.3.3 Flexible Benefits of Time Distribution in Practice

The present study aimed to examine the benefits of optimal spacing by measuring and comparing short and long intervals between sessions. In other words, it aimed to identify how much time is required by learners between sessions, in order for them to improve their vocabulary learning and retention. The results showed that both short and long spacing was sufficient for learners to progress. Therefore, a possible reason for both Experimental Groups progressing equally could be the spacing of sessions; whether this interval was one or seven

days, it was enough for them to retain vocabulary. This is linked to the previous literature, which reveals an advantage for spaced over massed practice, including the pilot study in this present research. There appears to be consistent evidence that allowing some space between practice sessions is beneficial. However, the research into exactly how long that spacing should be has produced mixed results. In this regard, there may be other factors coming into play, such as individual differences (Larsen-Freeman, 2014), as mentioned previously in subsection 6.3.1. It suggests that the cohort of learners with low English ability who were sampled for this particular study needed an interval of one or seven days between their practice sessions. It is an important result, because it gives both teachers and students flexibility in teaching and learning, both within and outside the confines of their institute.

Traditional teaching methods do not appear to be meeting the needs of this group of learners, because they were just not progressing prior to this study. They were placed at Level 1 (low-beginner) based on their results in the placement test (see subsection 3.4.1). The VLT tests and the pre-test, which were taken before the intervention, produced rather low scores for the current participants, reflecting their somewhat poor vocabulary knowledge (see subsection 4.4.1 and section 4.5). Therefore, the teaching method implemented, using this computer-assisted means of teaching vocabulary and evidenced by the spacing of practice sessions, was beneficial for this particular group of students in their vocabulary learning and retention. This method of vocabulary teaching offers educational institutes, teachers and students much greater flexibility in the timetabled teaching sessions, within a traditional classroom experience. It means that the delivery of the curriculum can be adapted to overcome other constraints that impact on teaching. For example, in intensive summer courses, teachers could adopt an intensive method of teaching vocabulary. Meanwhile, during term time (the Autumn and Spring Terms), the standard version of weekly-spaced delivery would be helpful. Thus, teachers would have the flexibility to provide teaching based on the constraints of their schools, schedules and the students' availability. This is because both Experimental Groups made strong gains, and the different spacing did not impact their results. It suggests that either short or long intervals between sessions can be beneficial, with teachers being afforded some flexibility based on their local constraints.

Aside from the above, if a school is closed for any reason, such as weather conditions, traditional teaching methods are impacted by the lost classroom time. However, CALL allows

flexibility, so that this lost time can be made up by the students, because it provides a high degree of independence; for example, in the form of access outside the classroom (Reinders & Hubbard, 2013). Thus, the students can be more autonomous, taking control of when and where they access learning resources and applications. Nevertheless, the risk is that they will avoid doing the work. However, both the Intensive and Spaced Groups' answers to the Quizlet questionnaire showed that Quizlet was motivating for them (see section 6.4), thereby indicating their intention to use it in the future. What it is not known is for how long the students might sustain their future use of Quizlet. Moreover, the study demonstrates that the technology appeared to be helpful for vocabulary learning, because the students in both Experimental Groups used it and did well. This is in line with Sanosi's (2018) study, which supports the use of Quizlet as a beneficial tool for learning vocabulary, whether inside or outside university (see subsection 2.4.2.4).

In fact, Quizlet is a freely available software that can help eliminate the constraints of a school timetable. However, the use of such technology could also introduce new problems. For example, classrooms in Oman are not all equipped with computers. Usually, a PC and a data display screen are available for the teacher in a regular classroom, and Level 1 students in the College are allowed to use computers in the school computer labs twice a week. The advantage is that the students have some classes in the labs, but the disadvantage is that there are many classes seeking to access the labs to use the computers. Therefore, it is important for the College to look at increasing students' access to computers, so that they can use them to practise between lessons, as some students may not have a computer at home. Moreover, this could also reduce the risk of students failing to practise at home.

The results of the present study are encouraging for technology use, because of the positive impact of Quizlet. Although Quizlet is available as a phone app, the students were not allowed to use their phones in this case, because the researcher wanted to control their exposure to the target vocabulary. Therefore, it is not known whether the results would have been different had the students been given unlimited access to the application. However, the use of mobile phones in the classroom is not permitted amongst students at this College, which is a policy to facilitate classroom management and avoid any disruption in class. The researcher also noted that few students on the Foundation Programme brought a laptop or tablet with them, except their own mobiles. This could limit their flexibility for

computer/mobile learning at the College. Therefore, it might be better in future studies to look at the use of mobile phones, in order to see if this could help with vocabulary learning, indicating that further research on mobile learning is required.

6.3.4 The Participants and the Technique Feature Analysis (TFA) Framework in Practice

As mentioned previously, the participants in this present study were part of a new intake of students placed in Level 1 (pre-elementary level), because of their low results in a placement test administered on their enrolment at the College (see subsection 3.4.2). However, almost all schools are likely to have a group of students who are not able to progress beyond the pre-elementary level of language learning, and it is difficult to demonstrate progress with such groups of learners. Families and the wider society may feel that it is essential to acquire English language skills (see subsection 6.2.2), but some learners seem not to be progressing in this area. With such external pressure, whether from families, society, or the schools themselves, a learner can reach the age of the adult students in this current study, feeling inwardly hopeless about ever improving their English skills (Lin, 2012; Rodriguez & Abreu, 2003). This is self-defeating, as these students will not be motivated to learn. Therefore, the findings of this current study are very important, showing how these young people successfully learned the target vocabulary, whereupon they made significant progress, despite their previous level of English language proficiency and educational history.

Nevertheless, the researcher acknowledges that there may be a number of possible factors influencing the equally significant progress demonstrated by both Experimental Groups. For example, it could have been due to the amount of rehearsal performed, which helped with the recall of information (Mukoyama, 2004) by both groups. This amount of rehearsal was not experienced by the Control Group, who were a test-only group. Therefore, the Control Group's results in the immediate post- and delayed post-tests for Recognition and Recall remained far below those of the two Experimental Groups. The Experimental Groups' high degree of progress reflects the attempts made in this present study, based on Baddeley and Hitch's (1974) Working Memory Model. These authors emphasise the importance of rehearsing information repeatedly to transfer information from the short-term to the long-term memory (Baddeley, 2000). This technique, along with mechanisms for linking L2 vocabulary

with its L1 meaning was found to support long-term vocabulary retention (Jiang, 2000). In addition, images, sounds, writing and spelling retrieval all helped the students to 'notice' and then recall information (Coxhead, 2008). These are tools that Quizlet makes available. Therefore, the present researcher believes Quizlet tools and activities play a major role in enabling the students to learn new vocabulary.

The Technique Feature Analysis (TFA), proposed by Nation and Webb (2011), was implemented to assess Quizlet vocabulary activities in this current study. Hu and Nassaji (2016) advise involving more features of the TFA Framework in any vocabulary learning activity. The finding showed that the adopted vocabulary teaching method in the current study using Quizlet matches approximately 78% of the features of the TFA Framework (see section 3.4.4.8 & Table 3.8). This finding highlighted that the Quizlet activities were of good quality. It has identified the key beneficial features of the Quizlet practice. The quality of this practice may offer one explanation for why no lag effect (i.e. benefit for longer spacing) was observed. In other words, it may be that the intensity of practice is less critical when the practice itself is high quality. However, this question requires further research. The analysis of Quizlet conducted using the TFA framework highlights the teaching quality of the Quizlet activities. It is worth mentioning here that Hu and Nassaji (2016, p.31) claimed: "No empirical studies, however, have yet examined the predictive power of TFA" in a real practice setting. Consequently, this result is considered to be an original contribution, supporting the TFA Framework as a powerful indicator for vocabulary-learning activities within a real practice setting with students of low English language ability.

6.3.5 Summary of Discussion for RQ2

In response to the second research question, the answer is affirmative: both time distributions of the practice sessions (Intensive versus Spaced) were equally sufficient for the Experimental Groups to benefit from using Quizlet in practice. Thus, neither of the Experimental Groups displayed any differences in terms of improving their vocabulary learning and retention. This finding is important, because when looking at the skills that must be acquired when learning a language (listening, speaking, reading and writing), learners need to be able to understand and recognise words that will help them take the initial steps towards understanding that language. This first language learning step (i.e. recognition) was not being achieved with this sample of

students when traditional teaching methods were used. Therefore, the present study evaluated the learners in their Recognition and Recall knowledge, as well as their Passive and Active knowledge in both Recognition and Recall. It was not only aimed at helping them improve their receptive knowledge by understanding the meanings of words, but also at helping them demonstrate this knowledge by recalling the form. The anticipated result was that the learners would make more progress in their basic English skills. The results for Recognition showed that both Experimental Groups maintained similarly high levels of progress in the 4-day RI and 28-day RI delayed post-tests. However, there appeared to be a drop in the Recall results for the Intensive Group in both delayed post-tests. This drop also occurred in the Spaced Group in the longer-term, delayed post-test. However, none of these drops in Recall were reliable, because of the very small effect sizes, with confidence intervals crossing zero.

Therefore, the present study suggests that either one-day or seven-day spacing can be beneficial as shorter or longer retrieval intervals. This gives schools, teachers and students some flexibility in the face of constraints on delivering vocabulary instruction. In addition, this study supports the use of a TFA Framework to evaluate the effectiveness of vocabulary activities. Consequently, this study could be considered as one of very few that supports the use of this framework in an authentic classroom. It also fills the gap in the literature by focusing on vocabulary learning using Quizlet and two different time distributions of practice sessions (Intensive versus Spaced), in order to help learners who were at the pre-elementary stage of English language-learning in a real-life setting. Therefore, the results of this intervention could explain and fill the knowledge gap relating to the use of the TFA Framework to measure the effectiveness of vocabulary learning activities by assessing learners' short- and long-term retention in practice. The present researcher believes that such teaching methods using Quizlet could play a major role in improving learners' English vocabulary knowledge.

6.4 RQ3. What Are the Perceptions of Teachers and Students on the Use and Implementation of Quizlet?

With the aim of investigating the participants' perceptions of using Quizlet for vocabulary learning, one-to-one interviews were conducted with four core teachers of the Experimental Groups (Intensive and Spaced) (see section 5.3). In addition, the Quizlet questionnaire

(including 25 closed-ended and three open-ended questionnaire items) was administered to 87 students from both Experimental Groups (see section 5.2). The independent samples t-test (parametric) and Mann-Whitney test (non-parametric) were run for the questionnaire's quantitative data, in order to identify any significant differences between the Experimental Groups (see section 5.2). Next, the findings derived from the quantitative (survey) and qualitative data (open-ended questions and the interviews) were analysed and organised under three essential themes of the Technology Acceptance Model (TAM), developed by Davis (1989) (see subsection 2.4.2.4). These themes include perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI). Turning now to RQ3, the overall results of the interviews and Quizlet questionnaire suggest that the participants (the students and their teachers) in both Experimental Groups held positive views of using Quizlet in terms of PEOU, PU and BI (see section 3.2). This will be discussed in the following subsections, with each theme being addressed individually.

6.4.1 Quizlet's Perceived Ease of Use (PEOU)

In terms of ease of use, the quantitative (survey) data notably indicated that the students from both Experimental Groups considered Quizlet to be uncomplicated software that was very easy to use. The Intensive and Spaced Groups recorded high mean scores of 4.07 and 3.80, respectively, for items under this theme (see subsection 5.2.1). However, on inspecting the students' responses, most but not all the participants declared Quizlet to be easy to use. Nevertheless, only a few mentioned experiencing some initial difficulties in using this software programme, especially when trying to open an account. This could reflect certain individual differences between students in their abilities and familiarity with using technology. It could also explain the significant difference between the Intensive and Spaced Groups when running the parametric and non-parametric tests (see subsection 5.2.1). However, these differences did not prevent either of the Experimental Groups from progressing equally and sustaining this progress.

Similarly, the teachers agreed that Quizlet was straightforward and uncomplicated for their students. This finding is contrary to some previous studies conducted in Oman (Al-Musawi, 2007; Al-Senaidi et al., 2009), as these earlier studies revealed negative views amongst teachers, with regard to the use of technology in the classroom. The teachers in these studies mentioned some of the reasons for their negative views, including their lack of

technology awareness (Al-Musawi, 2007) and limited technology support (Al-Senaidi et al., 2009). This lack of technology awareness and knowledge is identified as an essential domain in the Technology, Pedagogy, and Content Knowledge framework (TPACK) (Mishra & Koehler, 2006), which provides a useful schema for teachers seeking to integrate technology into the classroom. The TPACK model suggests that a combination of technological, pedagogical and content knowledge is required for technology to be used effectively in the classroom (Koehler & Mishra, 2009).

Therefore, the teachers in the present study emphasised the importance of providing both teachers and students with two to three training sessions in the use of the programme. This would suggest that it was essential to provide teachers with a number of training sessions, especially in light of the massive push towards integrating technology into higher education institutions (HEIs), not only in Oman, but also globally (Al Musawi, 2007). It points to the need to introduce any application or software programme before its implementation, so that awareness can be raised of its potentially positive impact on language learning, and to support staff in integrating technology effectively into their classrooms. Kim (2002) emphasises the importance of including teachers in any learning process, because their own positive perceptions of technology as teachers will in turn positively influence their students' knowledge of its implementation. Therefore, the current study gathered the perceptions of both students and teachers, who agreed that the programme was clear and easy to use. It should also be noted here that the teachers and students in this study were given a brief explanation of the programme prior to the intervention.

6.4.2 Quizlet's Perceived Usefulness (PU)

With regard to the usefulness of Quizlet for learning vocabulary, both the teachers and students found Quizlet beneficial. For the students, this result was reflected in the high mean scores obtained by the Intensive and Spaced Groups in their responses under the PU theme, amounting to 3.94 and 3.70, respectively (see subsection 5.2.2). These results concur with Al-Khatib's (2011) study, wherein it was found that the use of technology had a positive impact on learners' achievements in tests, and enhanced their understanding and ability to communicate. This positive influence of technology was also interpreted from the high degree of progress achieved by both Experimental Groups in the immediate post-test, with this progress being sustained in both the delayed (4-day RI; 28-day RI) post-tests.

Based on Rohrer and Pashler's (2007) optimal spacing of RI, which favours the Spaced Group (with 25% of RI) over the Intensive Group (with 3.5% of RI) at 28-day RI, it was predicted that the Intensive Group's results would be reduced at the delayed (28-day RI) post-test. However, the Intensive Group maintained their progress in the long term (at 28-day RI), demonstrating no statistically significant differences from the Spaced Group. This unexpected result for the Intensive Group could be due to the fact that they were highly engaged and motivated, a finding which was revealed by running an independent samples t-test and Mann-Whitney test for the PU theme items (see subsection 5.2.2). This analysis revealed that the Intensive Group expressed more positive views of using Quizlet than the Spaced Group. Their enthusiasm for Quizlet could therefore be a reason for their sustained progress at the delayed (28-day RI) post-test. This finding corresponds to Al-Khatib's (2011) observation of the positive influence of technology in enhancing learners' achievements in a final exam. However, it should be noted that the researcher did not use any additional motivation measurements for the students in the present study.

Both the teachers and learners listed a number of features in Quizlet that made it beneficial for vocabulary learning. First, most of the teachers and students found that L1 translation was useful for retaining words. This is in line with Jiang's (2000) model and Schmitt's (2008) views on using L1 translation to help learners build up an initial connection between form and meaning, as a way of facilitating L2 vocabulary acquisition. Although the teachers held different views of the usefulness of the Match game activity on Quizlet, all the teachers and students agreed that it was the most beneficial and motivating activity for the students. The teachers also agreed that Quizlet's Write produced the highest rate of word production. Therefore, although the teachers agreed that this Write activity could be difficult for the learners, they found it to be a very useful practice activity. Moreover, both the teachers and students agreed that the multimedia aids (images and sounds) played an important role in clarifying the meaning and pronunciation of words, which corresponds to the results of Zapata and Sagarra's (2007) study. Zapata and Sagarra found that most of the learners sampled in their study held positive views of using technology to enhance their pronunciation and vocabulary knowledge. This view is further supported by Nation (2013) and Farhangi et al. (2016), who claim that the direct, explicit association of words with images prevents any interference in meaning. In addition, Lin et al. (2008) revealed that EFL learners were able to improve their pronunciation, vocabulary knowledge, and speaking and reading skills in this

way. Therefore, the present study adds further evidence to the views presented in the existing literature, which support that technology use is beneficial for language learning.

6.4.3 The Participants' Behavioural Intention (BI)

The findings for the BI theme, derived from the interviews and questionnaires, revealed that the teachers and students were willing to utilise Quizlet in future to learn vocabulary. This willingness to continue using Quizlet, demonstrated by almost all the students and their teachers, gives a strong indication of their full satisfaction with the programme. Both the Intensive and Spaced Groups achieved very close mean scores of 3.32 and 3.21, respectively, for this point (see subsection 5.2.3). In addition, the Mann-Whitney test revealed no significant differences between the two Experimental Groups. Therefore, the Intensive and Spaced Groups appeared to agree that they would employ Quizlet for their future vocabulary learning.

The teachers also added that Quizlet could be a superior alternative to Moodle (an online platform used at the College). This was because they noticed that their students were much more engaged by Quizlet than by Moodle. All the teachers agreed that Quizlet motivated their students to learn vocabulary. These observations and positive comments from teachers, concerning the use of Quizlet, were supported in this present study by the significant improvement achieved by the students in both Experimental Groups, as indicated in the results of performance tests (pre-, immediate post- and delayed post-tests) (see section 4.5).

Overall, the results of this current study support previous research findings (Chien, 2015; Dizon, 2016). For example, Dizon (2016) emphasises the positive influence of CALL in general and Quizlet in particular, with regard to improving vocabulary knowledge in the classroom. In fact, this current study adapted Dizon's questionnaire and formulated the interview questions based on Davis's (1989) three main TAM themes (PEOU, PU and BI). Moreover, the findings of the present study align with those of Dizon's work, conducted on a very limited number of students (nine in total). Additionally, this positive view expressed by the students in the present research corresponds to the learners' views in Chien's (2015) study. However, Chien's study did not consider teachers' opinions on using Quizlet in the classroom. Indeed, previous studies on the use of Quizlet (Chien, 2015; Dizon, 2016; Jackson III, 2015) have mainly concentrated on students' perceptions. Therefore, the results of the

current study contribute to the existing knowledge by filling the gap in the literature on teachers' perceptions. Moreover, this study is unique in that it was conducted in a real-life setting, with a large number of participants (87 students; four teachers), using multiple research methods (questionnaires and interviews).

6.4.4 Summary of Discussion for RQ3

In response to the final research question, the answer is that both the teachers and students found Quizlet easy to use and useful. Thus, they were willing to use it in future. The researcher adopted Davis's (1989) Technology Acceptance Model (TAM) to organise and categorise the participants' positive views. Different research methods (questionnaires and interviews) were deployed to obtain deep and clear answers to the third research question in this study. Consequently, both the qualitative (interviews and open-ended questions) and quantitative (survey) findings support the overall perception of Quizlet as a user-friendly and uncomplicated software. The participants (teachers and students) also considered Quizlet to be valuable and beneficial for learning vocabulary. Therefore, they demonstrated an intention to use Quizlet in future.

CHAPTER SEVEN: CONCLUSION

7.1 Introduction

The aim of this last chapter is to conclude the current thesis. It therefore begins with a brief summary of the study and the main findings. The study's limitations and recommendations for further research are also suggested, followed by the pedagogical implications for the classroom. Finally, the contributions of the research are presented.

7.2 Summary of the Study

The present thesis reports on a quasi-experimental study, which attempted to evaluate the effectiveness of using Quizlet for vocabulary learning and retention under long (7-day) and short (1-day) spaced practice schedules, with low-ability learners in a classroom-based learning context. The participants were four English teachers and their 96 Level 1 students (aged 18 to 19 years), enrolled on the Foundation Programme at an Omani College of Technology. The participating students had low English language proficiency and continued to face difficulties in learning English, despite the fact that they had started learning English at an early age, from Year 1 in school. This thesis also investigated the perceptions of both sets of participants (the students and their teachers), with regard to using Quizlet for vocabulary learning.

The participants were drawn from six intact classes and divided into three groups: one Control Group (a test-only group) and two Experimental Groups (one Intensive with 1-day ISI and one Spaced, with 7-day ISI). All three groups completed baseline tests, comprising three vocabulary level tests (2,000, 3,000 and 5,000 VLT), a background questionnaire, and two WM tests (forward and backward digit span). The groups were then administered three performance tests (pre-, immediate and delayed). In addition, between the pre- and immediate post-test, the Experimental Groups (Intensive and Spaced) received an intervention treatment of four vocabulary practice sessions using Quizlet, followed in the final practice session by a questionnaire to elicit their perceptions of using Quizlet. Each practice session lasted for 20 minutes. The Intensive Group undertook their sessions on four consecutive days, while the Spaced Group received the intervention over the course of four weeks at a rate of one session per week. The Experimental Groups' core teachers subsequently participated in one-to-one

interviews with the researcher, so that their views on the effectiveness of using Quizlet could be gathered.

The researcher was aware of a potential influence that being the former colleague of the interviewees might have on responses. For example, the researcher could not ensure the neutrality in answering questions (Hesse-Biber & Leavy, 2010). Moreover, researchers might directly or indirectly ask leading questions that could provide answers that researchers want to hear, not revealing the actual opinions of interviewees (Berg, 2016). Therefore, it is important to confirm that participants were told the aims of the study, the purpose of the interview and the importance of understanding their beliefs and perceptions. Further, the researcher confirmed the importance of protecting and respecting their confidentiality and anonymity. Therefore, any indirect influence of the researcher on the qualitative responses obtained from the interviewees was minimised. Moreover, in the intervention, the researcher acted as a facilitator to guide learners in using Quizlet and ensure that all classes had as similar experience as possible within the research design to ensure an equal amount of time for each activity.

The Quizlet software programme was utilised in this study to practise 34 target words (nouns) during an intervention treatment for two Experimental Groups. These words were all taken from the Level 2 Vocabulary Log, consisting of either 2,000-level vocabulary or academic words. This list was selected to ensure that the words would be more advanced than the participants' current level. Quizlet contains a number of vocabulary activities. However, the researcher administered activities that were oriented towards vocabulary recognition and recall. The amount of time and types of vocabulary task were determined so that the researcher could give equal treatment to both Experimental Groups and control their exposure to the target words. These steps were performed to help reduce the effect of external factors, which had the potential to affect the study results.

The methodological design of the present thesis involved dividing each of the three main Groups (Intensive, Spaced and Control) into two sub-groups: the 4-day RI sub-groups, who took the delayed post-test four days after the last practice session, and the 28-day RI sub-groups, who took the delayed post-test 28 days after the last practice session. All results of the performance tests were then analysed according to this distribution.

7.3 Summary of the Findings

The analyses of the quantitative data (performance tests and survey) and qualitative data (open-ended questions and interviews) gave a substantial indication of the effectiveness of Quizlet's practical application in the classroom among young adult learners with low English ability. The participants in the Experimental Groups (Intensive and Spaced) showed significant progress in the immediate post-test, which they sustained in the short-term (at 4-day RI) and long-term (at 28-day RI), delayed post-tests. Moreover, the findings showed no significant differences between the Experimental Groups over the three performance tests: pre-, immediate post and delayed. However, the Control Group did not display any changes over the three performance tests, and their results were significantly lower than the results of both Experimental Groups in the immediate and delayed post-tests. The findings therefore reflect the effectiveness of Quizlet practice among this particular cohort of learners.

The significant progress made by the Experimental Groups in this study supports the effectiveness of explicit vocabulary learning for low-ability learners, as an ideal method of vocabulary learning and retention for this group. In addition, the decrease in the number of missing answers revealed the role of technology in general and Quizlet in particular, with regard to motivating learners and inspiring confidence in them to attempt more answers. Moreover, the number of fully correct answers (i.e. correct spelling and correct meaning) also generated an increase in the students' scores. This means that the learners were highly engaged with the activities provided by Quizlet. All these findings were further supported by the perceptions of both teachers and students, as they expressed positive views of using Quizlet, considering it to be an effective tool of vocabulary learning and retention. Therefore, this finding corroborates the hypothesis underpinning the first research question, namely that Quizlet is an effective method of learning vocabulary for low-ability learners, who are seeking to increase their vocabulary knowledge through practice.

When evaluating the short-term (1-day ISI) and long-term (7-day ISI) time distribution between the practice sessions (Intensive versus Spaced), the findings revealed no significant differences between the Experimental Groups across the three performance tests (pre-, immediate post- and delayed post-tests). Both groups showed equal improvement from Time 1 to Time 2, and from Time 1 to Time 3. They also maintained equal progress from Time 2 to

Time 3, with no significant differences. These findings violate the two hypotheses attached to the second research question, where optimal spacing was adopted (optimal ISI=between 10% and 30% of the RI range, as suggested by Rohrer & Pashler, 2007; 25% was investigated) (see subsection 6.3.1). The findings were interesting and unexpected, because they contradicted this optimal spacing, which gave preference to the Intensive Group in short-term retention (4-day RI) and the Spaced Group in long-term retention (28-day RI). Thus, along with the mixed findings from the previous literature on the lag effect, the present study suggests that there is no single, optimal spacing that suits all learners. This could be due to differences between individual learners, or because intervals of one and seven days were sufficient for most of the participants, or because the Quizlet activities were beneficial in themselves. The researcher used the TFA Framework, which includes five main criteria – motivation, noticing, retrieval, generation and retention – in order to assess the vocabulary activities on Quizlet. The total percentage match with the TFA Framework criteria was around 78%, which could provide a rationale for the effectiveness of using Quizlet as a beneficial tool for teaching vocabulary. This also supports the TFA Framework as a useful indicator to assess vocabulary-learning activities assigned to learners. In this thesis, the importance of spacing practice sessions for low-ability learners is considered, whether at short (1-day) or long (7-day) intervals. This allowed the teachers and learners some flexibility over the vocabulary teaching and learning, according to their timetabled teaching sessions and availability inside and outside the confines of their institute. In addition, spacing the sessions allowed the participants to rehearse vocabulary at each practice session, thereby enabling information to be transferred from the working memory into the long-term memory, based on Baddeley and Hitch's (1974) Working Memory Model. In addition, multimedia content (sound and images) is important for noticing and recalling information (Coxhead, 2008). Finally, the L1 translation provided by Quizlet could also play a role in transferring information into the long-term memory (Jiang, 2000). Therefore, it would seem that Quizlet is an effective software programme for learning vocabulary.

Considering the importance of teachers' and students' views on the teaching methods or tools used in the learning process, the researcher formulated the questions in the Quizlet questionnaire and interview, in conformity with the Technology Acceptance Model (TAM) and according to three themes: perceived ease of use (PEOU), perceived usefulness (PU) and behavioural intention (BI). The findings show that both the teachers and students generally

considered the programme to be easy and uncomplicated to use, valuing it as a beneficial and enjoyable vocabulary-learning tool. Consequently, they planned to use it in future.

7.4 Limitations of the Study and Recommendations for Future Research

As with any other study, this current project had a number of inevitable limitations, which need to be acknowledged before making recommendations for further research and indicating the pedagogical implications. The first limitation relates to the sampling of the participants and study design, in that the participants were not randomly selected. Random sampling could be considered more robust for this study design, suggesting a limitation in this research. However, the researcher needed to work within a real practice setting; within the classroom setting it was not possible to randomly assign students within each class to the three experimental groups, because of practical constraints with the classrooms and school timetable and due to the potential risk of cross-contamination between the two experimental and control groups. Although non-random sampling could be considered a limitation of an experimental design, this study was aiming to reflect the reality through investigating this application in real practice, which is one of the key contributions of this study to the knowledge. Therefore, a quasi-experimental design was adopted, in which two intact classes were assigned to each Experimental Group. It is important to note that there were inherent difficulties in conducting the intervention simultaneously across six classes, divided into three treatment conditions (Intensive, Spaced and Control). It was not possible to combine the two classes in each group into a single classroom, because of their different timetables, and the lack of a large enough classroom to accommodate each group. Nevertheless, having more than one class in each group could create a balance that would avoid any external factors or possible bias influencing the data, such as different levels of teaching engagement in each class.

Another limitation was associated with the performance tests (pre-, immediate and delayed) and pilot study. The data from the performance tests were gathered at three different time points. Thus, the test effect may have influenced the participants' results in the post-tests (immediate and delayed). To reduce this test effect, the researcher included a control group and manipulated the test items in different arrangements across the three performance tests. Although the test effect could be a potential issue, the Control Group's performance

suggested that any gains observed in the Experimental Groups were not due to the test effects, as the Control Group did not demonstrate learning gains at the post- and delayed post-tests.

It should be noted that before conducting the main study, a pilot study was carried out over a period of three weeks. The researcher acknowledges that the sample size and time period of the pilot study was very limited. However, its main purpose was to obtain more information about the research instruments and the difficulties that might be faced in actual practice (see subsection 3.4.5).

A further limitation of this study was the use of a Level 1 sample (pre-elementary level on the Foundation Programme). These participants had a relatively low level of English proficiency and all came from just one College. Consequently, the results may not be generalisable to other students with higher levels of English language proficiency, or to other institutions in Oman or elsewhere. Therefore, it could be advisable in future studies to conduct a similar study on students with a higher level of English ability in different types of educational institution. In addition, further research in a real classroom setting is required, as this simulates the reality of students in the classroom and reflects their authentic need to progress.

Finally, this present study implemented Quizlet without comparing it with other vocabulary-learning software. Although the participants were found to have positive views of using Quizlet and achieved highly significant scores, Quizlet cannot be claimed to be the most effective software tool for learning vocabulary, because no comparison was made between different types of learning software in this study. Therefore, further intervention studies comparing vocabulary-learning software programmes are recommended. In particular, further research into the use of mobile phones and their integration into English preparation classes would be worthwhile, in order to investigate their potential for facilitating vocabulary teaching and learning both within and outside of the classroom.

7.5 Pedagogical Implications in the Classroom

In a wider pedagogical context, the findings of the current research clearly indicate the pedagogical advantages of explicit vocabulary learning through a variety of activities provided via Quizlet for low-ability learners. This corroborates Nation and Meara (2010),

who support the use of explicit methods of teaching vocabulary to learners in their initial stages of language learning. Therefore, this thesis bears out the effectiveness of Quizlet in developing vocabulary knowledge amongst students with low English language ability in an Omani college. Based on this finding, teachers are encouraged to provide vocabulary instruction using technology as a vehicle for the multimedia content provided in such programmes, with a view to enhancing vocabulary knowledge and pronunciation (Lin et al., 2008; Zapata & Sagarra, 2007). In addition, this thesis proposes the integration of technology in general and Quizlet in particular into the classroom, as a means of teaching and learning vocabulary.

The findings also indicate that the students in the Experimental Groups (Intensive and Spaced) enjoyed using Quizlet as an opportunity to learn vocabulary. Therefore, this thesis suggests that teachers and the College enable better access to computers for this cohort of learners. For example, more computer labs are needed for students. The students could also be encouraged to download a Quizlet mobile-learning application to work on at home, if they do not have adequate access to computers at College, due to the high demand for computers in teaching.

Furthermore, this thesis highlights the importance of providing spacing between practice sessions to allow for the transfer of information into the long-term memory. The findings for the participants across the performance tests suggest that the low-ability learners in both Experimental Groups made significant progress, whether the practice was intensive or spaced. This means that the different spacing used in this present study did not impact on learners' results, suggesting that either 1-day ISI or 7-day ISI could be beneficial. Therefore, teachers can be flexible based on other constraints bearing upon their teaching, such as the length of the term or duration of summer courses (see subsection 6.3.3 for further details).

As mentioned previously, both teachers and students in the Experimental Groups favoured Quizlet to learn vocabulary. However, the teachers noted their lack of technological awareness, which can cause difficulties for using technology in the classroom. Therefore, the thesis emphasises the importance of raising the awareness and knowledge of technology use, before implementing any software programmes. In order to do so, the researcher suggests introducing Quizlet, or any other programme intended for integration into the classroom, to both teachers and students, before the actual implementation of that programme. It is essential

to involve both teachers and learners in any course planning, because taking their views into account will have a positive influence on the language-learning process (Kim, 2002).

Finally, the researcher suggests using the Technique Feature Analysis (TFA) Framework (Nation & Webb, 2011) to assess any vocabulary learning activities that might be used in the classroom. This thesis has provided evidence of the effectiveness of using the TFA Framework to evaluate vocabulary activities (see subsection 6.3.4). This suggestion is associated with Hu and Nassaji's (2016) recommendation to include more features of TFA in vocabulary-learning activities. Therefore, the use of the TFA Framework was essential in this study for assessing the teaching method adopted, namely the use of Quizlet in practice.

7.6 Contributions of the Study

To conclude, this present thesis makes a significant contribution to the field of vocabulary learning, using technology and time distribution. First, a mixed methods design was adopted to avoid potential methodological limitations. Quantitative analyses were carried out to investigate the impact of time distribution (i.e. spacing between practice sessions; intensive versus spaced) on vocabulary learning and retention in low ability learners of L2 English using a CALL programme (Quizlet). These were supplemented by semi-structured interviews with four core teachers of the Experimental Groups and employed the Quizlet questionnaire to the Experimental Groups, for the purpose of exploring the teachers' and students' perceptions on the use and implementation of Quizlet.

This thesis is significant for several reasons. First, the main contribution of the study to knowledge lies in its sampling of students with a low level of English language proficiency. Its second main contribution is its examination of a significant practical issue in a real-life setting, rather than in the laboratory. It consequently goes beyond a theoretical clinical design by adopting a research approach (i.e. quasi-experimental) that is normally associated with clinical studies; applying it instead in practice, specifically with lower-ability learners. Furthermore, only a limited number of empirical research studies have investigated the spacing effect (Pavlik & Anderson, 2005) on low-level learners and most of the existing research has compared massed with spaced distributions of presentation. Added to this, only a handful of studies have begun to look at lag effects in L2 vocabulary learning by manipulating the length of intervals between practice sessions, and exploring how the spacing

of practice sessions affects the knowledge acquired, and the length of time that this knowledge is retained (Serrano & Huang, 2018). Therefore, this study constitutes a novel contribution to knowledge by endeavouring to fill the gaps in the existing literature by bridging the gap between theory and practice through being practice-based and implementing the study in a real-life setting. In addition, the study adopted an explicit method of vocabulary-teaching via Quizlet, under two different spaced practice schedules to facilitate vocabulary-learning and evaluate students' long-term retention using two delayed post-tests (4-day RI and 28-RI).

Aside from the above, this study is the first of its kind to be carried out at an Omani College of Technology. However, although it was conducted in a single college in Oman, it was anticipated that its findings would be of interest to other countries, because understanding the impact of spacing practice sessions on vocabulary learning is of relevance to language classrooms in general. Therefore, it is considered that the present research will contribute to research on the impact of lag effects (longer versus shorter spacing) in vocabulary learning by measuring four types of vocabulary knowledge: receptive recall, productive recall, receptive recognition and productive recognition (Laufer & Goldstein, 2004).

The current study used Quizlet for vocabulary acquisition within a theoretical framework presented by Davis (1989) (see subsection 2.4.2.4). This was the first time that the TAM Framework had been used for effective vocabulary learning via available technology at the College. Most of the previous research (for example, Chien, 2015; Dizon, 2016; Jackson III, 2015) has focused on learners' attitudes and perceptions of using Quizlet, without attempting to control any variables in vocabulary learning, such as parts of speech, frequency, practice tasks, or measurement tools. The existing research is often limited to the opinions of learners outside the Omani context. Therefore, the results of this research will contribute to the literature on vocabulary acquisition, considering that Quizlet plays an essential role in motivating students to learn vocabulary (Jackson III, 2015). One of the original contributions to knowledge made by this study is therefore its identification of an effective instructional approach to vocabulary learning and teaching, in light of the lag effects observed. In addition, this study's original contribution is based on the fact that it is the first to assess Quizlet activities in actual practice, using the TFA Framework. Here, this framework is considered to be a powerful indicator for the effectiveness of vocabulary activities.

Finally, the gap in the literature on teachers' opinions was considered in this study, with the aim of helping to fill this gap in the existing knowledge. Kim (2002) emphasises the importance of teachers' positive views about using technology, which affect positively the implementation of technology in the classroom and the language-learning process. Thus, it is expected that the findings of this study will encourage reflection on how vocabulary is presented and practised (e.g. using Quizlet) and when vocabulary is practised and revisited within and outside of the classroom. A final point is that it is crucial for curriculum designers to present comprehensive explications of the benefits of implementing and integrating various vocabulary-teaching methods, including technology use. This is as important in Oman as it is anywhere else.

7.7 Summary

This Conclusion summarises the main points stated and explained in this thesis. It also presents the main study findings and considers the research limitations, before making suggestions for future research. In addition, the study's pedagogical implications in the classroom are outlined. Finally, this chapter concludes by emphasising the significant contributions made by this thesis to the fields of vocabulary teaching and learning.

References

Abdelgadir, E. M., & Ramana, V. L. (2016). Challenges of teaching English to Arabic students. *International Journal of English Language, Literature and Humanities, 4*(11), 221-227.

Afrin, N. (2014). Integrating computer assisted instruction in the EFL Classroom of Bangladesh. *IOSR Journal of Humanities and Social Science, 19*(11), 69-75.

Al Jabri, B. A. (2008). *Vocabulary teaching techniques in an Omani government school* (Unpublished Master's thesis). American University of Sharjah, Sharjah, UAE.

Al Rawahy, H. (2001). E-learning: A telecommunications perspective. In *Educational technology symposium/exhibition 2001 proceedings,* Sultan Qaboos University, Muscat, Sultanate of Oman.

Al Riyami, T. K. S. (2016). *Introducing critical pedagogy to English language teachers at Tertiary Education in Oman: Attitudes, potentialities and challenge: A critical postmodernist study* (Published doctoral dissertation). University of Exeter, UK.

AL-Hammadi, F. S. (2012). The role of recognition memory in L2 development. *Journal of King Saud University-Languages and Translation, 24*(2), 83-93.

Al-Issa, A. S. (2006). The cultural and economic politics of English language teaching in Sultanate of Oman. *Asian EFL Journal, 8*(1), 194-218.

Al-Issa, A. S., & Al-Bulushi, A. H. (2012). English language teaching reform in Sultanate of Oman: The case of theory and practice disparity. *Educational Research for Policy and Practice, 11*(2), 141-176.

Al-Jadidi, H. S. S. (2009). *Teaching English as a foreign language in Oman: An exploration of English language teaching pedagogy in tertiary education* (Unpublished doctoral dissertation). Victoria University, Melbourne, Australia.

Al-Jardani, K. S. (2012). English language curriculum evaluation in Oman. *International Journal of English Linguistics, 2*(5), 40-44.

Al-Khatib, H. (2011). Technology enhanced learning: Virtual realities; Concrete results case study on the impact of TEL on learning. *European Journal of Open, Distance and E-Learning*. *14*(1). 1–12.

Al-Mahrooqi, R., Al-Maamari, F., & Denman, C. (2016). An analysis of lexicon in Omani English language teaching: Challenges from teaching and learning perspectives. *Journal of Arts and Social Sciences, 7*(1), 27-48.

Al-Mamari, A. S. (2012, February). *General foundation program in higher education institutions in Oman national standards: Implementation and challenges*. Paper presented at the Quality & Enhancement in Higher Education, Oman Quality Network Regional Conference. Muscat, Sultant of Oman.

Al-Musawi, A. S. (2007). Current status of educational technologies at Omani higher education institutions and their future prospective. *Educational Technology Research and Development, 55*(4), 395-410.

Al-Musawi, A., & Abdelraheem, A. (2004). E-learning at Sultan Qaboos University: Status and future. *British Journal of Educational Technology, 35*(3), 363-367.

Al-Senaidi, F. S., & Wyatt, M. (2014). Involving female Omani English language teachers in evaluating curriculum materials. *Argentinian Journal of Applied Linguistics, 2*(2), 42-59.

Al-Senaidi, S., Lin, L., & Poirot, J. (2009). Barriers to adopting technology for teaching and learning in Oman. *Computers & Education, 53*(3), 575-590.

Allsop, Y. (2016). Does technology improve learning? The value of constructivist approaches to technology-based learning. *ICT in Practice*. http://www.ictinpractice.com/does-technology-improve-learning-the-value-of-constructivist-approaches-to-technology-based-learning/

Alsadoon, R. A. (2015). *Vowel blindness: Computer-mediated help options for Arabic EFL learners* (Unpublished doctoral dissertation). Simon Fraser University, British Columbia, Canada.

Ambu-Saidi, S. S. K. (2010). *Student and teacher attitudes towards the use of language learning technology (LLT) in a tertiary English course in Oman.* (Unpublished Master's thesis). The University of Queensland, Brisbane, Australia.

Anjaniputra, A. G., & Salsabila, V. A. (2018). The merits of Quizlet for vocabulary learning at tertiary level. *Indonesian EFL Journal, 4*(2), 1-11. https://doi: 10.25134/ieflj.v4i2.1370.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation* (vol. 2, pp. 89-195). Academic Press.

Attamimi, R. A. (2018). Pronunciation encumbrances for Omani EFL students at the English

language centre of the Salalah College of Technology: A pre observational study. *International Journal of English Language and Linguistics Research, 6*(3), 7-10.

Averianova, I. (2015). Vocabulary acquisition in L2: Does CALL really help?. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 30-35). Research-publishing.net. http://dx.doi.org/10.14705/rpnet.2015.000306

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education, 5*(6), 66-70.

Baddeley, A. D. (1978). The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review, 85*(3), 139-152.

Baddeley, A. D. (1986). *Working memory*. Oxford University Press.

Baddeley, A. D. (1990). *Working memory*. Lawrence Erlbaum Associates.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory?. *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of learning and motivation* (vol. 8, pp. 47-89). Academic Press.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long-and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*(2), 176-189.

Bagheri, E., Roohani, A., & Ansari, D. N. (2012). Effect of CALL-based and non-CALL based methods of teaching on L2 vocabulary learning. *Journal of Language Teaching and Research, 3*(4), 744-752.

Barr, B. W. B. (2016). Checking the effectiveness of Quizlet as a tool for vocabulary learning. *The Center of EFL Journal, 1*(2), 36–48. https://doi.org/10.15045/ELF_0020104.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge University Press.

Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography, 6*(4), 253-279.

Beatty, K. (2013). *Teaching & researching: Computer-assisted language learning*. Routledge.

Berg, B. L. (2016). *Qualitative research methods for the social sciences* (9th ed.). Pearson Education.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*(1), 44-62.

Bird, S. (2010). Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics, 31*(4), 635–650.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Brown, G. D., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed?. *Journal of Memory and Language, 34*(5), 594-621.

Bryman, A. (2015). *Social research methods*. Oxford University Press.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*(3), 551-581.

Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review, 24*(3), 369–378.

Cellat, S. (2008). *Computer assisted vocabulary learning: A study with Turkish 4th grade EFL learners* (Unpublished Master's thesis). Anadolu University, Eskişehir, Turkey.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-380.

Chien, C. W. (2015). Analysis the effectiveness of three online vocabulary flashcard websites on L2 learners' level of lexical knowledge. *English Language Teaching, 8*(5), 111-121.

Childers, J. B., & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. Hirsh-Pasek, & R. Golinkoff (Eds.), *Action meets word: How children learn* verbs (pp. 311-355). Oxford University Press.

Chiswick, B. R., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development, 26*(1), 1-11.

Climie, E. A., & Rostad, K. (2011). Test review: Wechsler Adult Intelligence Scale (4th ed.). *Journal of Psychoeducational Assessment, 29*(6), 581-586.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3–23). Ablex Publishing Corporation.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education* (5th ed.). Routledge.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge.

Cohen, L., Manion, L., & Morrison, K. (2011). Surveys, longitudinal, cross-sectional and trend studies. In L. Cohen, L. Manion, & K. Morrison (Eds.), *Research methods in education* (7th ed., pp. 256-289). Routledge.

Cook, V., & Bassetti, B. (2005). An introduction to researching second language writing systems. In V. J. Cook, & B. Bassetti (Eds.), *Second language writing systems* (pp. 1-67). Multilingual Matters.

Courtney, L. (2014). *Moving from primary to secondary education: An investigation into the effect of primary to secondary transition on motivation for language learning and foreign language proficiency* (Unpublished doctoral dissertation). University of Southampton, Southampton, UK.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of working memory*: *Mechanisms of active maintance and active control* (pp. 62–101). Cambridge University Press.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in Brain Research, 169,* 323-338.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Coxhead, A. (2008). Phraseology and English for academic purposes. In F. Meunier, & S. Granger (Eds.), *Phraseology in language learning and teaching* (pp. 149-161). John Benjamins Publishing.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Craik, F. I., & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience & Biobehavioral Reviews, 36*(7), 1729-1739.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294.

Crandell, E. R. (2017). *Quizlet flashcard for the first 500 words of the Academic Vocabulary List* (Unpublished Master's thesis). Brigham Young University, Provo, Utah.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage Publications.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Crotty, M. (2005). *The foundations of social research: Meaning and perspectives in the research process.* Sage Publications.

Davies, G. (2000). CALL (computer assisted language learning). In M. Byram (Ed.), *Routledge encyclopedia of language teaching and learning* (pp. 90-93). Routledge.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly,13*(3), 319-340.

Davis, N. (2017). *The selfish gene*. Macat Library.

Denscombe, M. (2014). *The good research guide: For small-scale social research projects*. (5th ed.). Open University Press.

Dewey, J. (1938). *Experience and education*. Simon & Schuster.

Dizon, G. (2016). Quizlet in the EFL classroom: Enhancing academic vocabulary acquisition of Japanese university students. *Teaching English with Technology, 16*(2), 40-56.

Doolittle, P. E. (2007). *Constructivism and online education*. In Proceedings Of online Conference on teaching online in higher education, Fort Wayne, USA.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first-and second-language learners. *Reading Research Quarterly*, *38*(1), 78-103.

Drubin, D. G., & Kellogg, D. R. (2012). English as the universal language of science: Opportunities and challenges. *Molecular Biology of the Cell, 23*(8), 1399. https:// doi: 10.1091/mbc.E12-02-0108

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 73–87). Law- rence Erlbaum.

Egbert, J. (2005). *CALL essentials: Principles and practice in CALL classrooms.* TESOL Publications.

ELCS (The English Language Curriculum Section). (2010). *English language curriculum framework*. Ministry of Education: Sultanate of Oman.

Ellis, N. C. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. *Computer Assisted Language Learning, 8*(2-3), 103-128.

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33-68). Cambridge University Press.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*(2), 305-352.

Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language acquisition research.* Routledge.

Elsayyad, H. M. (2014). *The relationship between working memory and reading comprehension in L1 Arabic and L2 English for Arabic speaking children* (Unpublished doctoral dissertation). Bath Spa University, Bath, England.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19-23.

Eysenck, M. W. (1978). Levels of processing: A reply to Lockhart and Craik. *British Journal of Psychology, 69*(2), 177-178.

Farhangi, B., Mohseni, S., & Ahmadi, M. (2016) Involvement load hypothesis and technique feature analysis. *Language Teaching and Testing, 1*(4), 184-188.

Fender, M. (2008). Spelling knowledge and reading development: Insights from Arab ESL learners. *Reading in a Foreign Language, 20*(1), 19–42.

Fernández, F. (Ed.). (2001). *Los estudios ingleses en el umbral del tercer milenio*. Servei de Publicacions de la Universitat de València.

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). Sage Publications Ltd.

Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition, 26*(2), 349-356.

Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: A first–person approach. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 356*(1413), 1351-1361.

Goossens, N. A., Camp, G., Verkoeijen, P. P., Tabbers, H. K., & Zwaan, R. A. (2012). Spreading the words: A spacing effect in vocabulary learning. *Journal of Cognitive Psychology, 24*(8), 965-971.

Graves, K., & Xu, S. (2000). *Designing language courses: A guide for teachers (No. 428 G7.).* Heinle & Heinle Publishers.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 371-377.

Greene, R. L. (1992). *Human memory: Paradigms and paradoxes.* Psychology Press.

Greenhouse, SW, & Geisser, S.(1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112.

Guba, E. G. (Ed.). (1990). *The paradigm dialog*. Sage publications.

Hammersley, M. (2007). *Educational research and evidence-based practiceh.* Sage Publications Ltd.

Hani, N. A. B. (2014). Benefits and barriers of computer assisted language learning and teaching in the Arab world: Jordan as a model. *Theory and Practice in Language Studies*, *4*(8), 1609-1615.

Hasman, M. A. (2004). The role of English in the 21st century. *TESOL Chile, 1*(1), 18-21.

Hatch, M. J. with Cunliffe, AL (2006) *Organization theory: Modern, symbolic and postmodern perspectives* (2nd ed.). Oxford University Press.

Heatley, A., Nation, I.S.P. & Coxhead, A. (2002). *RANGE and FREQUENCY programs*. http://www.victoria.ac.nz/lals/staff/paul-nation.aspx

Hemchua, S., & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect, 21*(3), 3–25.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(02), 303-317.

Hesse-Biber, S. N., & Leavy, P. (2010). *The practice of qualitative research* (2nd ed.). Sage publications.

Higgins, S., Xiao, Z., & Katsipataki, M. (2012). *The impact of digital technology on learning: A summary for the education endowment foundation.* Education Endowment Foundation and Durham University.

Hu, H. C. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement load hypothesis versus technique feature analysis. *System*, *56*(1), 28-39.

Huberman, M., & Miles, M. B. (2002). *The qualitative researcher's companion.* Sage Publications.

Hulstijn, J.H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language, 17*(1), 23-59.

Hunt, M., & Feng, J. (2016). *Improving vocabulary of English language learners through direct vocabulary instruction.* Presentation at Chinese American Educational Research and Development Association Annual Conference. Washington, USA.

Huyen, N. T. T., & Nga, K. T. T. (2003). Learning vocabulary through games. *Asian EFL Journal, 5*(4), 90-105.

Ismail, A. (2011). *Language planning in Oman: Evaluating linguistic and sociolinguistic fallacies* (Unpublished doctoral dissertation). Newcastle University, Newcastle, UK.

Jackson III, D. B. (2015). A targeted role for L1 in L2 vocabulary acquisition with mobile learning technology. *Perspectives (TESOL Arabia), 23*(1), 6-11.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*(1), 47–77.

Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*(4), 617-637.

Johnson, B., & Christensen, L. (2008). *Educational research: Quantitative, qualitative, and mixed approaches* (3rd ed.). Sage Publications.

Józsa, K., & Morgan, G. A. (2017). Reversed items in Likert scales: Filtering out invalid responders. *Journal of Psychological and Educational Research, 25*(1), 7-25.

Kamanpoori, H. (2014, May 5). Making 12 years of effective English learning at schools is important. *Oman Observer*. http://omanobserver.om/making-12-years-of-effective-english-learning-at-schools-is-important/

Kasprowicz, R. E., Marsden, E., & Sephton, N. (2019). Investigating distribution of practice effects for the learning of foreign language verb morphology in the young learner classroom. *The Modern Language Journal*, *103*(3), 580-606.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365-386.

Kilickaya, F., & Krajka, J. (2010). Comparative usefulness of online and traditional vocabulary learning. *TOJET: The Turkish Online Journal of Educational Technology, 9*(2), 55-63.

Kim, H. (2002). Teachers as a barrier to technology-integrated language teaching. *English Teaching*, *57*(2), 35-64.

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, *58*(2), 285-325.

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)?. *Contemporary Issues in Technology and Teacher Education*, *9*(1), 60-70.

Korlu, H., & Mede, E. (2018). Autonomy in vocabulary learning of Turkish EFL learners. *The EuroCALL Review*, *26*(2), 58-70.

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, *23*(9), 1297-1317.

Krish, P., Hussin, S., & Sivapuniam, N. (2011). Language learning and language acquisition in online forums. *3L: Language, Linguistics, Literature*, *17*(2), 91-100.

Kuiper, K. (2009). *Formulaic genres.* Palgrave Macmillan.

Küpper-Tetzel, C. E., Erdfelder, E., & Dickhäuser, O. (2014). The lag effect in secondary school classrooms: Enhancing students' memory for vocabulary. *Instructional Science*, *42*(3), 373-388.

Kvale, S. (2008). *Doing interviews.* Sage Publications.

Kwiatkowska, G. (2007). *Techniques in teaching vocabulary at the beginners level.* http://www.profesor.pl/mat/pd7/pd7_g_kwiatkowska_20070420.pdf

Lander, B. (2016). Quizlet: What the students think–a qualitative data analysis. In S. Papadima- Sophocleous, L. Bradley, & S. Thouësny (Eds), *CALL communities and culture – short papers from EUROCALL 2016* (pp. 254-259). Research-publishing.net. https://doi.org/10.14705/rpnet.2016.eurocall2016.571

Larsen-Freeman, D. (2014). *Complexity Theory: Renewing Our Understanding of Language, Learning, and Teaching*. Keynote presentation at International TESOL Convention. Portland, Oregan. http://www.tesol.org/attend-and-learn/international-

Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS.* Routledge.

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

Laufer, B. (2006). Comparing focus on form and focus on forms in second-language vocabulary learning. *Canadian Modern Language Review, 63*(1), 149-166.

Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399-436.

Lee, L. (2005). Using web-based instruction to promote active learning: Learners' perspectives. *CALICO Journal, 23*(1), 139-156.

Lees, D. (2015). A brief comparison of digital-and self-made word cards for vocabulary learning. *Kwansei Gakuin University Humanities Review, 18,* 59-71.

Levy, M. (1997). *Computer assisted language learning*. Clarendon Press.

Lin, L. C. (2012). *Measuring adult learners' foreign language anxiety, motivational factors, and achievement expectations: A compa- rative study between Chinese as a second-language students and english as a second-language students* (Unpublished doctoral dissertation). Cleveland State University, Cleveland, Ohio.

Lin, L. F. (2010). English learners' incidental vocabulary acquisition in the video-based CALL program. *The Asian EFL Journal, 12*(4), 51-66.

Lin, S., Winaitham, W., & Saitakham, K. (2008). *The use of websites for practicing listening skills of undergraduate students.* A case study at Suranaree University of Technology, Thailand. https://files.eric.ed.gov/fulltext/ED500929.pdf

Lincolin, Y. S., & Guba, E. (1985). *Naturalistic inquiry.* Sage publications.

Lotfolahi, A. R., & Salehi, H. (2016). Learners' perceptions of the effectiveness of spaced learning schedule in L2 vocabulary learning. *SAGE Open, 6*(2), 1-9.

Lotfolahi, A. R., & Salehi, H. (2017). Spacing effects in vocabulary learning: Young EFL learners in focus. *Cogent Education, 4*(1), 1-10.

Louis, C., Lawrence, M., & Keith, M. (2007). *Research methods in education.* Routledge.

Lyons, A. C. (2015). Approaches to collecting data. In P. Rohleder, & A. C. Lyons (Eds), *Qualitative research in clinical and health psychology* (pp. 74-91). Palgrave MacMillan.

Ma, Q. (2013). Computer assisted vocabulary learning: Framework and tracking users' data. In P. Hubbard, M. Schulze, & B. Smith (Eds), *Learner – computer interaction in*

language education (pp. 230-243). Computer Assisted Language Instruction Consortium (CALICO) Monograph.

Ma, Q., & Kelly, P. (2006). Computer assisted vocabulary learning: Design and evaluation. *Computer Assisted Language Learning, 19*(1), 15-45.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design.* Lawerence Erlbawn Associates.

Mahu, D. P. (2012). Why is learning English so beneficial nowadays?. *International Journal of Communication Research, 2*(4), 374-376.

Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Sage publications.

Marttinen, M. (2008). *Vocabulary learning strategies used by upper secondary school students studying English as a second language* (Unpublished Master's thesis). University of Jyväskylä, Jyväskylä, Finland.

Maskor, Z. M., & Baharudin, H. (2016). Receptive vocabulary knowledge or productive vocabulary knowledge in writing skill, which one important?. *International Journal of Academic Research in Business and Social Sciences, 6*(11), 261-271.

Masrai, A., & Milton, J. (2016). How many words do you need to speak Arabic? An Arabic vocabulary size test. *The Language Learning Journal* (a head of print). https://dx.doi.org/10.1080/09571736.2016.1258720

McLeod, S. A. (2008). Working memory. *Simply Psychology.* https://www.simplypsychology.org/working%20memory.html

McLeod, S. A. (2009). Short term memory. *Simply Psychology.* https://www.simplypsychology.org/short-term-memory.html

McLeod, S. A. (2010). Long-term memory. *Simply Psychology*. https://www.simplypsychology.org/long-term-memory.html

McLeod, S. A. (2019). Constructivism as a theory for teaching and learning. *Simply Psychology.* https://www.simplypsychology.org/constructivism.html

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412-433.

Mehring, J. G. (2005). *Developing vocabulary in second language acquisition: From theories to the classroom.*

http://www.hpu.edu/CHSS/LangLing/TESOL/ProfessionalDevelopment/200680TWPfall06/03Mehring. pdf.

Miles, S., & Kwon, C. J. (2008). Benefits of using CALL vocabulary programs to provide systematic word recycling, *English Teaching, 63*(1), 199-216.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.

Mishra, P., & Koehler, M.J. (2006). Technological pedagogical content knowledge: A framework for integrating technology in teacher knowledge. *Teachers College Record, 108*(6), 1017-1054.

Moras, S. (2001). *Computer-assisted language learning (CALL) and the Internet.* Karen's Linguistic Issues. http://www3.telus.net/linguisticsissues/CALL.html.

Mukoyama, M. (2004). *Efficiently memorizing words In foreign language acquisition* (Unpublished doctoral dissertation). Notre Dame Seishin University, Okayama, Japan.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*(3), 251-269.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*(1), 53-81.

Nakata, T. (2008). English vocabulary learning with word lists, word cards and computers: Implications from cognitive psychology research for optimal spaced learning. *ReCALL, 20*(1), 3-20.

Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: A critical investigation of flashcard software. *Computer Assisted Language Learning, 24*(1), 17-38.

Nakata, T. (2015). Effects of expanding and equal spacing on second language vocabulary learning: Does gradually increasing spacing increase vocabulary learning?. *Studies in Second Language Acquisition, 37*(4), 677-711.

Nakata, T., & Webb, S. (2016). Does studying vocabulary in smaller sets increase learning?. *Studies in Second Language Acquisition, 38*(3), 523-552.

Nasr, R. T. (1963). *The teaching of English to Arab students*. Longman.

Nation, , I. S. P. (1974). Techniques for teaching vocabulary. *English Teaching Forum,12*(3), 18-21.

Nation, , I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (vol. 14, pp. 6-19). Cambridge University Press.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House Publishers.

Nation, I. S. P. (1994). *New ways in teaching vocabulary. New ways in TESOL series: Innovative classroom techniques.* TESOL Incorporated Company.

Nation, I. S. P. (2001) *Learning vocabulary in another language.* Cambridge University Press

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening?. *Canadian Modern Language Review, 63*(1), 59-82.

Nation, I. S. P. (2013). *Learning vocabulary in another language.* Cambridge University Press.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9-13.

Nation, I. S. P., & Gu, P. Y. (2007). *Focus on vocabulary.* Macquarie University/NCELTR Publications.

Nation, I. S. P., & Meara, P. (2010) Vocabulary. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed., pp. 34-52). Edward Arnold.

Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary.* Heinle.

Neath, I., & Surprenant, A. M. (2003). *Human memory: An introduction to research, data, and theory* (2nd ed.). Wadsworth.

Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior, 16*(2), 151-171.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Norouzian, R., & Plonsky, L. (2018). Eta-and partial eta-squared in L2 research: A cautionary review and guide to more appropriate usage. *Second Language Research, 34*(2), 257-271.

Norton, P., & Wiburg, K. M. (2003). *Teaching with technology: Designing opportunities to learn*. Wadsworth/Thomson Learning.

Nunan, D. (1999). *Second language teaching & learning*. Heinle & Heinle Publishers.

Özer, Y. E., & Koçoğlu, Z. (2017). The use of Quizlet flashcard software and its effects on vocabulary learning. *Language Journal, 168*(1), 61-81.

Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science, 29*(4), 559-586.

Piaget, J. (1977). *The development of thought: Equilibrium of cognitive structure*. Viking Press.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912.

Porter, A. (2016). A helping hand with language learning: Teaching French vocabulary with gesture. *The Language Learning Journal, 44*(2), 236-256.

Pring, R. (2015). *The philosophy of education* (3rd ed.). Bloomsbury Publishing PLC.

Punch, K. (2009). *Introduction to research methods in education*. Sage Publications.

Punch, K. F., & Oancea, A. (2014). *Introduction to research methods in education.* Sage Publications..

Quizlet. (2018). *Our mission* | Quizlet. Retrieved March 28, 2018 from https://quizlet.com/mission

Rädle, P. (2009). VTrain (Vocabulary Trainer)–Awards. *VTrain. net.* http://www.vtrain.net/aw.htm

Rasekh, Z. E., & Ranjbary, R. (2003). Metacognitive strategy training for vocabulary learning. *Tesl-Ej, 7*(2), 1-15.

Reinders, H., & Hubbard, P. (2013). CALL and learner autonomy: Affordances and constraints. In H. Reinders & M. Thomas (Eds.), *Contemporary computer assisted language learning* (pp. 359-375). Bloomsbury Publishing.

Ritchie, G. (2016). *Semantic, phonological and episodic representations in verbal immediate serial recall* (Unpublished doctoral dissertation). Australian Catholic University. Melbourne, Australia.

Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R.(Eds.). (2013). *Qualitative research practice: A guide for social science students and researchers.* Sage Publications.

Rodriguez, M., & Abreu, O. (2003). The stability of general foreign language classroom anxiety across English and French. *The Modern Language Journal, 87*(3), 366- 374.

Rogers, J. (2015). Learning second language syntax under massed and distributed conditions. *TESOL Quarterly, 49*(4), 857–866.

Rogers, J. (2017). The spacing effect and its relevance to second language acquisition. *Applied Linguistics*, *38*(6), 906-911.

Rogers, J., & Cheung, A. (2018). Input spacing and the learning of L2 vocabulary in a classroom context. *Language Teaching Research*. Advance Online Publication. https://doi: 10.1177/1362168818805251.

Rohrer, D. (2015). Student instruction should be distributed over long time periods. *Educational Psychology Review*, *27*(4), 635-643.

Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Current Directions in Psychological Science*, *16*(4), 183-186.

Rovai, A., Baker, J., & Ponton, M. (2012). *Social science research design and statistics: A practitioner's guide to research methods and SPSS analysis*. Watertree Press.

Rovai, A., Baker, J. & Ponton, M. (2014). *Social science research design and statistics: A practioner's guide to research methods and IBM SPSS analysis* (2nd ed.). Watertree Press LLC.

Royse, D. (2004). *Research methods in social work.* Thomson Brooks/Cole.

Saigh, K., & Schmitt, N. (2012). Difficulties with vocabulary word form: The case of Arabic ESL learners. *System, 40*(1), 24-36.

Sanosi, A. B. (2018). The Effect of Quizlet on vocabulary acquisition. *Asian Journal of Education and e-Learning*, *6*(4) 71-77.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy.* Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N. (2005) *Vocabulary in language teaching* (6th ed.). Cambridge University Press.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing, 18*(1), 55-88.

Scott, M. (2008). *WordSmith Tools. Version 5.* Lexical Analysis Software. http://www.lexically.net/wordsmith/version5/index.html

SCT (Salalah College of Technology). (2018). *Facilities and Resources*. https://www.sct.edu.om/web/index.php/en/facilities-and-resources

Sergon, V. (2011). Playing the blame game: English education in Omani government schools. Independent Study Project (ISP) Collection. *Paper, 1132,*1-36. http://digitalcollections.sit.edu/cgi/viewcontent.cgi?article=2132&context=isp_collection

Serrano, R. (2012). Is time concentration good or bad for learning? Reflecting on the results from the cognitive psychology and from the SLA literature. In C. Muñoz (Ed.), *Intensive exposure experiences in second language learning* (pp. 3-22). Multilingual Matters.

Serrano, R., & Huang, H. Y. (2018). Learning vocabulary through assisted repeated reading: How much time should there be between repetitions of the same text?. *TESOL Quarterly*. *52*(4), 971-994.

Sharifi, M., Azizifar, A., Jamalinesari, A., & Gowhary, H. (2015). The Effect of Rosetta Stone computer software on vocabulary learning of Iranian elementary EFL learners. *Procedia-Social and Behavioral Sciences, 192,* 260-266.

Sharndama, E. C., Samaila, Y., & Tsojon, Y. I. (2014). English for Academic Purpose: A tool for enhancing students' proficiency in English language skills. *International Journal of English Language Teaching*, *1*(2), 14-20.

Sharoff, S. (2006). Creating general-purpose corpora using automated search engine queries. In M. Baroni, & S. Bernardini (Eds.), *WaCky! Working papers on the Web as Corpus.* Gedit.

Smith, C. B. (2003). Vocabulary development: Elaboration for writing. *ERIC Topical Bibliography and Commentary.* https://pdfs.semanticscholar.org/da1f/d86aad3e8698ab66af5e751f55efb2bf2af8.pdf

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139-152.

Statistics.laerd.com. (2019). *Independent T-Test - An introduction to when to use this test and what are the variables required* | Laerd Statistics. https://statistics.laerd.com/statistical-guides/independent-t-test-statistical-guide.php

Stephens, D. (2013). Frameworks for the sharing of memory in a digital world [abstract]. *Working Draft.*

https://scholar.google.co.uk/citations?user=9jVd5u4AAAAJ&hl=sl#d=gs_md_cita-d&u=%2Fcitations%3Fview_op%3Dview_citation%26hl%3Dsl%26user%3D9jVd5u4AAAAJ%26citation_for_view%3D9jVd5u4AAAAJ%3ATyk-4Ss8FVUC%26tzom%3D0

Sun, Y. C. (2003). Extensive reading online: An overview and evaluation. *Journal of Computer Assisted Learning, 19*(4), 438-446.

Sunderman, G. & Kroll, J.F. (2006). First language activation during second language lexical processing. *Studies in Second Language Acquisition, 28*(3), 387-422.

Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning*, *67*(3), 512–545.

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*(4), 860-895.

Suzuki, Y., & DeKeyser, R. (2017). Effects of distributed practice on the proceduralization of morphology. *Language Teaching Research*, *21*(2), 166-188.

Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning* (pp. 495-508). Routledge.

Swan, M. (1997). The influence of the mother tongue on second language vocabulary acquisition and use. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy.* Cambridge University Press.

Swehla, S. E., Burns, M. K., Zaslofsky, A. F., Hall, M. S., Varma, S., & Volpe, R. J. (2016). Examining the use of spacing effect to increase the efficiency of incremental rehearsal. *Psychology in the Schools, 53*(4), 404-415.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273-1296.

Tam, S. S., Kan, N. H., & Ng, L. L. (2010). Low proficiency learners in synchronous computer-assisted and face-to-face interactions. *Turkish Online Journal of Educational Technology-TOJET, 9*(3), 61-75.

Tang, C., & Treffers-Daller, J. (2016). Assessing incidental vocabulary learning by Chinese EFL learners: A Test of the involvement load hypothesis. In G. Yu, & Y. Jin (Eds.), *Assessing Chinese learners of English* (pp. 121-149). Palgrave.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage Publications.

Thompson, S. K. (2012). *Sampling* (3rd ed.). Wiley.

Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. In *Psychology of learning and motivation* (vol. 60, pp. 113-189). Academic Press.

Torut, B. (2000). Computer-assisted language learning: An overview. *Silpakorn University International Journal, 1*(1), 131-153.

Tsubaki, M. (2006). *Psychological accounts of foreign language vocabulary acquisition.* https://www.u-bunkyo.ac.jp/center/library/image/fsell2005_205-217.pdf

Tsubaki, M. (2007). *The involvement load hypothesis: An inquiry into vocabulary learning.* https://www.u-bunkyo.ac.jp/center/library/image/EIBUN175_184.pdf

Vallar, G., & Papagno, C. (1995). Neuropsychological impairments of short-term memory. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of memory disorders* (1st ed., pp. 135-165). Wiley.

Van Benthuysen, R. (2003). Explicit vocabulary instruction: Using a word list to focus attention. *Journal of Bunkyo Gakuin University, 2,* 89-96.

von Glasersfeld, Ernest (1984) An introduction to radical constructivism. In Watzlawick, P. (Ed) *The invented reality* (pp. 17-40). Norton.

von Glasersfeld, E. (1990). Environment and communication. In L. P. Steffe & T. Wood (Eds.), *Transforming children's mathematics education: International perspectives* (pp. 30-38). Lawrence Erlbaum Associates.

Vygotsky, L., S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University press.

Wallace, M. (1982). *Practical language teaching: Teaching vocabulary.* Heinemann Educational Books.

Wang, D. (2000). Vocabulary acquisition: Implicit learning and explicit teaching. In *REACT* (pp. 15-22). National Institute of Education. https://repository.nie.edu.sg/bitstream/10497/3825/3/REACT-2000-2-15.pdf

Webb, S. (2005). Receptive and productive vocabulary learning: The effect of reading and writing on word knowledge, *Studies in Second Language Acquisition 27*(1), 33-52.

Wei, R., Hu Y, and Xiong, J. (2019). Effect size reporting practices in applied linguistics: A study of one major journal. *SAGE Open, 9*(2), 1–19.

Wei, W. S., & Attan, A. (2013). Exploring strategies for vocabulary learning and teaching for ESL learners – A literature review. *Sains Humanika, 65*(2), 1-6.

Wellington, J. (2000). *Education research: Contemporary issues and practical approaches.* Continuum.

Wong, A. (2017). *The effect of degrees of reminders on vocabulary transfer to writing tasks* (Unpublished doctoral dissertation). Temple University, Philadelphia, Pennsylvania.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage Publications.

Yunus, M. M., Salehi, H., & Amini, M. (2016). Impact of using CALL on Iranian EFL learners' vocabulary knowledge. *English Language Teaching*, *9*(1), 173-187.

Zaki, H., & Ellis, R. (1999). Learning vocabulary through interacting with a written text. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 153-169). John Benjamins.

Zambrano Acosta, E. J. (2018). *(2018). Quizlet app and the English language vocabulary learning* (Unpublished Master's thesis). Universidad Técnica de Ambato. Ambato, Ecuador.

Zapata, G., & Sagarra, N. (2007). CALL on hold: The delayed benefits of an online workbook on L2 vocabulary learning. *Computer Assisted Language Learning*, *20*(2), 153-171.

Appendix 1 Ethical Approval to Conduct the Study

University of Reading
Institute of Education
Ethical Approval Form A (version May 2015)

University of Reading

Tick one:

Staff project: PhD_ ✓ EdD ____

Name of applicant(s): Muna Muqibal

Title of project: Effectiveness of CALL in Teaching and Learning English Vocabulary: The Case of Oman

Name of supervisor (for student projects): Dr Louise Courtney; Professor Cathy Tissot

Please complete the form below including relevant sections overleaf.

	YES	NO	
Have you prepared an Information Sheet for participants and/or their parents/carers that:			
a) explains the purpose(s) of the project	✓		
b) explains how they have been selected as potential participants	✓		
c) gives a full, fair and clear account of what will be asked of them and how the information that they provide will be used	✓		
d) makes clear that participation in the project is voluntary	✓		
e) explains the arrangements to allow participants to withdraw at any stage if they wish	✓		
f) explains the arrangements to ensure the confidentiality of any material collected during the project, including secure arrangements for its storage, retention and disposal	✓		
g) explains the arrangements for publishing the research results and, if confidentiality might be affected, for obtaining written consent for this	✓		
h) explains the arrangements for providing participants with the research results if they wish to have them	✓		
i) gives the name and designation of the member of staff with responsibility for the project together with contact details, including email. If any of the project investigators are students at the IoE, then this information must be included and their name provided	✓		
k) explains, where applicable, the arrangements for expenses and other payments to be made to the participants		✓	
j) includes a standard statement indicating the process of ethical review at the University undergone by the project, as follows: 'This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct'.	✓		
k) includes a standard statement regarding insurance: "The University has the appropriate insurances in place. Full details are available on request".	✓		
Please answer the following questions			
1) Will you provide participants involved in your research with all the information necessary to ensure that they are fully informed and not in any way deceived or misled as to the purpose(s) and nature of the research? (Please use the subheadings used in the example information sheets on blackboard to ensure this).	✓		
2) Will you seek written or other formal consent from all participants, if they are able to provide it, in addition to (1)?	✓		
3) Is there any risk that participants may experience physical or psychological distress in taking part in your research?		✓	
4) Have you taken the online training modules in data protection and information security (which can be found here: http://www.reading.ac.uk/internal/imps/Staffpages/imps-training.aspx)?	✓		
5) Have you read the Health and Safety booklet (available on Blackboard) and completed a Risk Assessment Form to be included with this ethics application?	✓		
6) Does your research comply with the University's Code of Good Practice in Research?	✓		
	YES	NO	N.A.
7) If your research is taking place in a school, have you prepared an information sheet and consent form to gain the permission in writing of the head teacher or other relevant supervisory professional?	✓		
8) Has the data collector obtained satisfactory DBS clearance?			✓
9) If your research involves working with children under the age of 16 (or those whose special educational needs mean they are unable to give informed consent), have you prepared an information sheet and consent form for parents/carers to seek permission in writing, or to give parents/carers the opportunity to decline consent?			✓

10) If your research involves processing sensitive personal data[1], or if it involves audio/video recordings, have you obtained the explicit consent of participants/parents?	✓		
11) If you are using a data processor to subcontract any part of your research, have you got a written contract with that contractor which (a) specifies that the contractor is required to act only on your instructions, and (b) provides for appropriate technical and organisational security measures to protect the data?			✓
12a) Does your research involve data collection outside the UK?	✓		
12b) If the answer to question 12a is "yes", does your research comply with the legal and ethical requirements for doing research in that country?	✓		
13a) Does your research involve collecting data in a language other than English?	✓		
13b) If the answer to question 13a is "yes", please confirm that information sheets, consent forms, and research instruments, where appropriate, have been directly translated from the English versions submitted with this application.	✓		
14a. Does the proposed research involve children under the age of 5?		✓	
14b. If the answer to question 14a is "yes": My Head of School (or authorised Head of Department) has given details of the proposed research to the University's insurance officer, and the research will not proceed until I have confirmation that insurance cover is in place.			✓
If you have answered YES to Question 3, please complete Section B below			

Please complete **either** Section A **or** Section B and provide the details required in support of your application. Sign the form (Section C) then submit it with all relevant attachments (e.g. information sheets, consent forms, tests, questionnaires, interview schedules) to the Institute's Ethics Committee for consideration. Any missing information will result in the form being returned to you.

A: My research goes beyond the 'accepted custom and practice of teaching' but I consider that this project has **no** significant ethical implications. (Please tick the box)	✓

Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category e.g. teachers, parents, pupils, etc.

The participants of this study will be Level 1 students from the Foundation Programme and their core English teachers

- For the pilot study: 1 teachers and 30 students
- For the main study: 4 teachers and 100 students

Give a brief description of the aims and the methods (participants, instruments and procedures) of the project in up to 200 words noting:

1. title of project
2. purpose of project and its academic rationale
3. brief description of methods and measurements
4. participants: recruitment methods, number, age, gender, exclusion/inclusion criteria
5. consent and participant information arrangements, debriefing (attach forms where necessary)
6. a clear and concise statement of the ethical considerations raised by the project and how you intend to deal with then.
7. estimated start date and duration of project.

Effectiveness of CALL in Teaching and Learning English Vocabulary: The Case of Oman.

This study aims to explore the impact of using Quizlet in vocabulary learning amongst students with early stages of English proficiency, in order to compare the effectiveness of massed versus spaced distribution instruction for vocabulary retention, and to measure the perceptions of both teachers and students, regarding the use of Quizlet.

Different instruments will be used in this study to answer the research questions. Three main methods will be employed: tests, questionnaires, and interviews, in addition to the software programme, Quizlet, which is employed as a learning method. It is worth mentioning that the attached student consent form and the questionnaire are in English, however, the Arabic versions for both documents will be provided for students.

The study will implement a quasi-experimental method, where four classes at Level 1 of the Foundation Programme will be allocated to two groups (massed and spaced), using a simple random sampling technique. The sample will consist of around 100 students (males and females), aged between 18 and 19 years old.

[1] Sensitive personal data consists of information relating to the racial or ethnic origin of a data subject, their political opinions, religious beliefs, trade union membership, sexual life, physical or mental health or condition, or criminal offences or record.

In the intervention, they will use a Quizlet study set created by the researcher to learn 15 verbs and 15 nouns through several activities provided by Quizlet using this link (https://quizlet.com/_40rl4t). The intervention will last for 4 weeks one session per week for 15 min for the spaced group, while all the activities will be given to the massed group at the same session for 60 min. Then, both groups will have an immediate test and two weeks later they will conduct a delayed post-test.

The activities are a part of standard teaching provision for the students, but the participant in this study is volunteering, with the right to withdraw without repercussions at any time during the project.

After the intervention, students will be asked to complete an attitude questionnaire about using Quizlet for vocabulary learning. Teachers will be asked to attend semi-structured interviews to find out their opinions about the effectiveness of using Quizlet in teaching vocabulary. Teachers will be interviewed individually for approximately 30 minutes at college during a time convenient to them. The interviews will be recorded and transcribed with teachers' permissions. The transcriptions will then be shown to teachers, in order to check their accuracy and confirm that they are still happy for their comments to be used.

It is important for the researcher to identify participants using their academic ID in the questionnaire to compare their results in the post-tests and their opinions towards using Quizlet. In case there are any extreme values in the results of some participants, the researcher can easily refer to the questionnaire to identify their background knowledge, gender or attitudes. In order to protect the anonymity of each participant, pseudonyms will be used to ensure participants cannot be identified.

There are no risks associated with taking part in this study. The information given by the participants will be kept strictly confidential and will only be seen by the researcher. It will not be possible to identify the staff, students or college in any published report resulting from this study.

The estimated start date of the pilot study is during the 2nd term, directly after receiving ethical approval. This process is likely to take four weeks. The main study will be conducted in the Autumn Term, 2018.

B: I consider that this project **may** have ethical implications that should be brought before the Institute's Ethics Committee.	

Please state the total number of participants that will be involved in the project and give a breakdown of how many there are in each category e.g. teachers, parents, pupils etc.

Give a brief description of the aims and the methods (participants, instruments and procedures) of the project in up to 200 words.

1. title of project
2. purpose of project and its academic rationale
3. brief description of methods and measurements
4. participants: recruitment methods, number, age, gender, exclusion/inclusion criteria
5. consent and participant information arrangements, debriefing (attach forms where necessary)
6. a clear and concise statement of the ethical considerations raised by the project and how you intend to deal with then.
7. estimated start date and duration of project

C: SIGNATURE OF APPLICANT:

Note: a signature is required. Typed names are not acceptable.

I have declared all relevant information regarding my proposed project and confirm that ethical good practice will be followed within the project.

Signed: …… Print Name: Muna Muqaibal Date…….

STATEMENT OF ETHICAL APPROVAL FOR PROPOSALS SUBMITTED TO THE INSTITUTE ETHICS COMMITTEE

This project has been considered using agreed Institute procedures and is now approved.

Signed: …… Print Name Prof Jill Porter Date…22/1/18….
(IoE Research Ethics Committee representative)*

* A decision to allow a project to proceed is not an expert assessment of its content or of the possible risks involved in the investigation, nor does it detract in any way from the ultimate responsibility which students/investigators must themselves have for these matters. Approval is granted on the basis of the information declared by the applicant.

Appendix 2 College Dean Consent

1st Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: c.tissot@reading.ac.uk
2nd Supervisor: Dr Rowena Kasprowicz
Tel.: +44 (0)118 378 2766
Email: r.kasprowicz@reading.ac.uk

Researcher: Muna Muqaibal

College Dean Information Sheet

Research Project:	Evaluating Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary the Classroom: The Case of Oman.
Research and Supervisors:	Muna Muqaibal (researcher); Professor Cathy Tissot; Dr Rowena Kasprowicz (supervisors)

Dear Sir,

I am writing to invite the English Language Center (ELC) at your college to take part in a research study about learning vocabulary.

What is this study?

This research forms the basis of a PhD project, which I am undertaking at the Institute of Education, University of Reading in the UK. It aims to explore the impact of using Quizlet in vocabulary learning amongst students with early stages level of English proficiency. This will compare the effectiveness of intensive versus spaced distribution instruction for vocabulary retention, and measure teachers' and students' perceptions of using Quizlet. It hopes to make recommendations regarding how we can best help teachers and students in these areas.

Why has this Center been chosen to take part?

I selected the ELC as the site of my research, because I have worked as a lecturer at this Center in the past. I am therefore familiar with the College's regulations and work environment. Consequently, I feel that the ELC, which offers the Foundation Programme (FP) to all students at the College, would be the best place for me to conduct my study amongst Level One students. In addition, the Center is equipped with language labs, which are important for my study, as I am using a software programme for learning vocabulary.

Does the Center have to take part?

There is no obligation for the Center to participate. This is entirely voluntary and the participants have the right to withdraw at any time by contacting Muna Muqaibal (the researcher) on Tel.: 99081666, email: m.h.a.muqaiball@pgr.reading.ac.uk. There will be no repercussions resulting from withdrawal for them or for anyone else concerned.

What will happen if the Center take part?

Subject to your agreement, participation in this study will involve administering an English vocabulary learning activity, using a software flashcard programme called Quizlet amongst Level One learners. These learners will be taught by the researcher and observed by their core classroom teachers on four occasions over four weeks. Before coming under treatment conditions, the participants will receive a background questionnaire, digit span test, grammar test, and vocabulary level test, in order to obtain an overview of their language skills and vocabulary knowledge. The digit span test will be conducted individually, with the participants in a separate room to measure their working memories and minimise disruption. Following the end of the treatment conditions, attitude questionnaires, and immediate and delayed post-tests will be administered. In addition, the core teachers will be interviewed individually and audio-recorded. Vocabulary learning activities will also be undertaken in normal lesson time in the ELC

1st Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: c.tissot@reading.ac.uk
2nd Supervisor: Dr Rowena Kasprowicz
Tel.: +44 (0)118 378 2766
Email: r.kasprowicz@reading.ac.uk

Researcher: Muna Muqaibal

language labs. The results gained from these tests will not be included in the students' grades. The recordings of all the interviews with participants will be transcribed and anonymised prior to data analysis.

What are the risks and benefits of taking part?

There are no risks associated with taking part in this study. The information provided by the participants will be kept strictly confidential and will only be seen by the researcher. It will not be possible to identify you, the teachers, the students, or the Center in any published report resulting from this study and information about individuals will not be shared with the Center.

I anticipate that the findings from this study will be useful for helping students to improve their vocabulary knowledge and for teachers in planning their vocabulary teaching.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. Research records will be stored securely in a locked filing cabinet and on a password-protected computer. Only the researcher will then have access to these records. No identifiers linking you, the participants, or the Center to this study will be included in any sort of report that might be published. The participants will be assigned a number and this will be used to refer to them in all records. All interview recordings will be destroyed after the end of the research. My academic supervisors will have access to the transcripts and test results, but I will be the only person accessing the original recordings. In line with the University's policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. The results of the study will be presented at national and international conferences, and in written reports and articles. We can send you electronic copies of these publications if you wish.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after the data collection has been completed, we will discard all data gathered at the Center.

What happens if something goes wrong?

In the unlikely case of any concerns or complaints, you can contact my supervisor, Dr. Louise Courtney at the University of Reading; Tel: 0118 378 2674; email: c.tissot@reading.ac.uk

Where can I get more information?

If you would like more information, please contact Muna Muqaibal.

I sincerely hope that you will give your consent for the Center to participate in this study. Should this be the case, please complete the attached consent form and return it to the researcher, Muna Muqaibal at the above email address.

1st Supervisor: Dr. Louise Courtney
Tel.: +44 (0) 118 378 2635
Email: l.m.courtney@reading.ac.uk
2nd Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: **c.tissot@reading.ac.uk**

Researcher: Muna Muqaibal

This project has been reviewed following the procedures of the University of Reading's Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

Thank you for your time.

Yours sincerely,

Muna Muqaibal

1st Supervisor: Dr. Louise Courtney
Tel.: +44 (0) 118 378 2635
Email: l.m.courtney@reading.ac.uk
2nd Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: **c.tissot@reading.ac.uk**

Researcher: Muna Muqaibal

University of Reading

College Dean Consent Form

I have read the Information Sheet about the project and received a copy of it.

I understand what the purpose of the project is and what is required of me. All my questions have been answered.

Name of College Dr. Hamdan AL Mantheri

Name of the college: [redacted] College of Technology

Please tick as appropriate:

I consent to the involvement of my college in the project as outlined in the Information Sheet ☐

Signed

Date: 19.2.2018

Appendix 3 Head of English Language Centre Consent

1st Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: c.tissot@reading.ac.uk
2nd Supervisor: Dr Rowena Kasprowicz
Tel.: +44 (0)118 378 2766
Email: r.kasprowicz@reading.ac.uk

Researcher: Muna Muqaibal

Head of English Language Center Information Sheet

Research Project:	Evaluating Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary the Classroom: The Case of Oman.
Researcher and Supervisors:	Muna Muqaibal (researcher); Professor Cathy Tissot; Dr Rowena Kasprowicz (supervisors)

Dear Sir,

I am writing to invite the English Language Center (ELC) to take part in a research study about learning vocabulary.

What is this study?

This research forms the basis of a PhD project, which I am undertaking at the Institute of Education, University of Reading in the UK. It aims to explore the impact of using Quizlet in vocabulary learning amongst students with early stages of English proficiency. This will compare the effectiveness of intensive versus spaced distribution instruction for vocabulary retention, and measure teachers' and students' perceptions of using Quizlet. It hopes to make recommendations regarding how we can best help teachers and students in these areas.

Why has this Center been chosen to take part?

I selected the ELC as the site of my research, because I have worked as a lecturer at this Center in the past. I am therefore familiar with the College's regulations and work environment. Consequently, I feel that the ELC, which offers the Foundation Programme (FP) to all students at the College, would be the best place for me to conduct my study amongst Level One students. In addition, the Center is equipped with language labs, which are important for my study, as I am using a software programme for learning vocabulary.

Does the Center have to take part?

There is no obligation for the Center to participate. This is entirely voluntary and the participants have the right to withdraw at any time by contacting Muna Muqaibal (the researcher) on Tel.: 99081666, email: m.h.a.muqaiball@pgr.reading.ac.uk. There will be no repercussions resulting from withdrawal for them or for anyone else concerned.

What will happen if the Center take part?

Subject to your agreement, participation in this study will involve administering an English vocabulary learning activity, using a software flashcard programme called Quizlet amongst Level One learners. These learners will be taught by the researcher and observed by their core classroom teachers on four occasions over four weeks. Before coming under treatment conditions, the participants will receive a background questionnaire, digit span test, grammar test, and vocabulary level test, in order to obtain an overview of their language skills and vocabulary knowledge. The digit span test will be conducted individually, with the participants in a separate room to measure their working memories and minimise disruption. Following the end of the treatment conditions, attitude questionnaires, and immediate and delayed post-tests will be administered. In addition, the core teachers will be interviewed individually and audio-recorded. Vocabulary learning activities will also be undertaken in normal lesson time in the ELC language labs. The

1st Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: c.tissot@reading.ac.uk
2nd Supervisor: Dr Rowena Kasprowicz
Tel.: +44 (0)118 378 2766
Email: r.kasprowicz@reading.ac.uk

Researcher: Muna Muqaibal

results gained from these tests will not be included in the students' grades. The recordings of all the interviews with participants will be transcribed and anonymised prior to data analysis.

What are the risks and benefits of taking part?

There are no risks associated with taking part in this study. The information provided by the participants will be kept strictly confidential and will only be seen by the researcher. It will not be possible to identify you, the teachers, the students, or the Center in any published report resulting from this study and information about individuals will not be shared with the Center.

I anticipate that the findings from this study will be useful for helping students to improve their vocabulary knowledge and for teachers in planning their vocabulary teaching.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. Research records will be stored securely in a locked filing cabinet and on a password-protected computer. Only the researcher will then have access to these records. No identifiers linking you, the participants, or the Center to this study will be included in any sort of report that might be published. The participants will be assigned a number and this will be used to refer to them in all records. All interview recordings will be destroyed after the end of the research. My academic supervisors will have access to the transcripts and test results, but I will be the only person accessing the original recordings. In line with the University's policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. The results of the study will be presented at national and international conferences, and in written reports and articles. We can send you electronic copies of these publications if you wish.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after the data collection has been completed, we will discard all data gathered at the Center.

What happens if something goes wrong?

In the unlikely case of any concerns or complaints, you can contact my supervisor, Dr. Louise Courtney at the University of Reading; Tel: 0118 378 2674; email: c.tissot@reading.ac.uk

Where can I get more information?

If you would like more information, please contact Muna Muqaibal.

I sincerely hope that you will give your consent for the Center to participate in this study. Should this be the case, please complete the attached consent form and return it to the researcher, Muna Muqaibal at the above email address.

1st Supervisor: Dr. Louise Courtney
Tel.: +44 (0) 118 378 2635
Email: l.m.courtney@reading.ac.uk
2nd Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: **c.tissot@reading.ac.uk**

Researcher: Muna Muqaibal

This project has been reviewed following the procedures of the University of Reading's Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

Thank you for your time.

Yours sincerely,

Muna Muqaibal

1st Supervisor: Dr. Louise Courtney
Tel.: +44 (0) 118 378 2635
Email: l.m.courtney@reading.ac.uk
2nd Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: **c.tissot@reading.ac.uk**

University of Reading

Researcher: Muna Muqaibal

Head of English Language Center (ELC) Consent Form

I have read the Information Sheet about the project and received a copy of it.

I understand what the purpose of the project is and what is required of me. All my questions have been answered.

Name of Head of ELC: SAEED SALIM ALMASHIKI

Name of College: [redacted] COLLEGE OF TECHNOLOG

Please tick as appropriate:

I consent to the involvement of my Center in this project, as outlined in the Information Sheet ☑

Signed

Date: 12. 02. 2018

Appendix 4 Teacher Consent

Researcher: Muna Muqaibal

1st Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: c.tissot@reading.ac.uk
2nd Supervisor: Dr Rowena Kasprowicz
Tel.: +44 (0)118 378 2766
Email: r.kasprowicz@reading.ac.uk

INFORMATION SHEET

Research Project: Evaluating Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary the Classroom: The Case of Oman.

Dear Teacher,

I am a PhD candidate at the University of Reading, UK. As part of the data collection stage of my thesis, I am writing to invite you to take part in a research study about vocabulary learning.

What is the study?

You have been invited to take part in this project, because I am looking to explore the impact of using Quizlet in vocabulary learning amongst students with early stages of English proficiency. This will enable a comparison to be made between the effectiveness of intensive and spaced distribution instruction for vocabulary retention. Moreover, it will help measure teachers' and students' perceptions of using Quizlet. It is hoped that recommendations can then be made, regarding how we can best help teachers and students in these areas. A total of approximately 154 potential participants have been invited to take part in this study, including 150 students and 4 teachers.

You are also invited, because you are a Level One lecturer at the English Language Center (ELC), given that my study applies to this level.

Do I have to take part?

It is entirely up to you whether you participate. You may also withdraw your consent to participation at any time during the project by contacting me, the Project Researcher, Muna Muqaibal, Tel.: 99081666, email: m.h.a.muqaiball@pgr.reading.ac.uk. This will not incur any repercussions to you.

What will happen if I take part?

If you agree to participate in this study, you will be asked to attend an interview with the researcher, lasting approximately 30 minutes. The interview will be recorded and transcribed with your permission. The transcription will then be shown to you, in order for you to check its accuracy and confirm that you are still happy for its contents to be used. The information gathered will be used by the researcher for data analysis.

What are the risks and benefits of taking part?

The information given by the participants in the study will remain confidential and will only be seen by the researcher and the supervisor listed at the top of this letter. You will be assigned an identification number (ID) and this will be the only reference used to distinguish your responses from those of other participants. This ID will in no way be associated with your name. The

records of this study will be kept private. No identifiers linking you to the study will be included in any sort of report that might be published.

I anticipate that the findings of this study will be useful for teachers, when planning how they teach vocabulary.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The data collected for this study will be kept private. No identifiers linking you, the students or the College to the study will be included in any sort of report that might be published. The participants will be assigned a number and this will be used to refer to them in all records. The research records will be stored securely in a locked filing cabinet and on a password-protected computer. Only the researcher and supervisors will have access to them. In line with the University's policy on the management of research data, anonymised data gathered in this research may be preserved and made publicly available for others to consult and re-use. The results of this study will be presented at national and international conferences, and included in written reports and articles. We can send you electronic copies of these publications if you wish.

What happens if I change my mind?

You can change your mind at any time, without any repercussions. If you change your mind after the data collection is complete, we will discard all the data.

What happens if something goes wrong?

In the unlikely event of any concerns or complaints, you can contact my supervisor, Dr. Louise Courtney at the University of Reading, Tel.: 0118 378 2674; email: c.tissot@reading.ac.uk

Where can I get more information about this study?

If you would like more information, please contact Muna Muqaibal.

I sincerely hope that you will agree to participate in this study. If you do, please complete the attached Consent Form and return it to me, Muna Muqaibal.

This project has been reviewed following the procedures of the University of Reading's Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

Signed:

Muna Muqaibal

Dated:

University of Reading

Consent Form

Project title: Evaluating Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary the Classroom: The Case of Oman.

I have read and had explained to me the Information Sheet relating to this project.

I have had explained to me the purposes of the project and what will be required of me. All my questions have been answered to my satisfaction. I agree to the arrangements described in the Information Sheet, insofar as they relate to my participation.

I understand that I will be interviewed and that the interview will be audio-recorded and transcribed.

I understand that my participation is entirely voluntary and that I have the right to withdraw from this project at any time, without giving a reason and without repercussions.

I have received a copy of this Consent Form and of the accompanying Information Sheet.

Please tick as appropriate:

I consent to being interviewed: _____ Yes _____ No

I consent to this interview being audio-recorded: _____ Yes _____ No

Name:

Signed:

Appendix 5 Student Consent (English version)

Researcher: Muna Muqaibal

1st Supervisor: Dr. Louise Courtney
Tel.: +44 (0) 118 378 2635
Email: l.m.courtney@reading.ac.uk
2nd Supervisor: Professor Cathy Tissot
Tel.: +44 (0) 118 378 2674
Email: **c.tissot@reading.ac.uk**

Students Information Sheet

The title: Evaluating Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary the Classroom: The Case of Oman.

Dear Participant,

I am a PhD student at the University of Reading, UK. You are invited to participate in this research study about vocabulary learning.

Please, take the time to read the following information carefully to know more about the research.

What is the study?

This study aims to explore the impact of using Quizlet in vocabulary learning amongst students with early stages of English proficiency. This will enable a comparison to be made between the effectiveness of intensive and spaced distribution instruction for vocabulary retention.

Why have I been chosen to take part?

You are invited, because you are a Level One student at the English Language Center (ELC). Your participation in the intervention and your views about Quizlet will help the researcher to find out the useful way to teach and learn vocabulary.

Do I have to take part?

Your participation is voluntary. Taking part will not influence your college grades in any way and information will not be shared with individual teachers. If you decide to take part, you will be given this information sheet to keep and asked to sign a consent form. You will be still free to withdraw at any time and without giving any reason by contacting me, the Project Researcher, Muna Muqaibal, Tel.: 99081666, email: m.h.a.muqaiball@pgr.reading.ac.uk.

What will happen if I take part?

The participation in this study will involve administering an English vocabulary learning activity, using a software programme called Quizlet. You will conduct the activities as a part of your standard teaching provision, but your permission is for sharing your data with the researcher. You will be also invited to complete a short questionnaire. It will take a maximum of 10 minutes to complete. The information gathered will be used by the researcher for data analysis.

What are the possible advantages and disadvantages of taking part?

You will find it useful to reflect on using Quizlet and help the researcher to find out the useful way of vocabulary learning and teaching.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The completed questionnaires of this study will be kept private. The data collected in the study will provide the basis of my PhD thesis. The thesis will be published in hard copy and electronic format which will be housed at the Institute of Education in the University of Reading. The data and the analysis of the data will also be used to produce articles, books, conference papers, as well as presented in conferences and lectures. In any of these formats I reassure you that the identity and anonymity of all participants will be protected. All information collected will be kept strictly confidential (subject to legal limitations). In order to protect the anonymity of each participant, pseudonyms will be used to ensure participants cannot be identified. All electronic data will be held securely in password-protected files on a non-shared PC and all paper documentation will be held in locked cabinets in a locked office.

What happens if I change my mind?

You can change your mind at any time, without any problems. If you change your mind after the data collection is complete, we will discard all the data.

What happens if something goes wrong?

In the unlikely event of any concerns or complaints, you can contact my supervisor, Dr. Louise Courtney at the University of Reading, Tel.: 0118 378 2635; email: l.m.courtney@reading.ac.uk

Where can I get more information about this study?

If you would like more information, please contact Muna Muqaibal.

I sincerely hope that you will agree to participate in this study. If you do, please complete the attached Consent Form and return it to me, Muna Muqaibal.

This project has been reviewed following the procedures of the University of Reading's Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

Signed:

Muna Muqaibal

Dated:

University of Reading

Student Consent Form

Project title: Evaluating Effectiveness of Spaced Practice using Computer-assisted Language Learning (CALL) in Teaching and Learning English Vocabulary the Classroom: The Case of Oman.

I have read the information sheet about the project and received a copy of it. I understand what the purpose of the study is, and what is required of me. All my questions have been answered.

Name of participant: ……

Please tick as appropriate:

1. I agree to participate in the intervention: _____ _____
   Yes No

2. I agree to take a part in the questionnaire: _____ _____
   Yes No

Name: _____ _____

Signed: _____ ____

Name of researcher taking consent: Muna Muqaibal

Researcher e-mail address:

Appendix 6 Student Consent (Arabic version)

Researcher: Muna Muqaibal

1st Supervisor: Professor Cathy Tissot
Email: c.tissot@reading.ac.uk
2nd Supervisor: Dr Rowena Kasprowicz
Email: r.kasprowicz@reading.ac.uk

نموذج صفحة المعلومات الخاصة بالدراسة

للطالب/ـة

عنوان الدراسة: تقييم فعالية تعليم اللغة بنظام التكرار المتباعد باستخدام الحاسوب في تدريس وتعلم المفردات الإنجليزية في الحصص الدراسي: دراسة حالة في سلطنة عمان.

عزيزي الطالب/ عزيزتي الطالبة،

انا طالبة دكتوراه بكلية التربية في جامعة ريدنج البريطانية. احتاج إلى مشاركة الطلاب والطالبات العمانيين في هذه الدراسة البحثية التي تهتم بتعلم المفردات الإنجليزية.
من فضلك اقرأ المعلومات التالية بعناية لمعرفة المزيد عن البحث.

ما هو الهدف من الدراسة؟

تهدف الدراسة إلى استطلاع آراء الطلبة عن تأثير استخدام برنامج كيزلت (Quizlet)، في تعلم المفردات الإنجليزية بين الطلبة في المستوى المبتدئ في اللغة الإنجليزية، ومقارنة فعالية التعليم بإجراء تمارين متعددة في أوقات متقاربة وأخرى في أوقات متباعدة لحفظ المفردات الإنجليزية.

لماذا يتم اختيارنا للمشاركة في الدراسة؟

بسبب دراستكم في المستوى الأول في البرنامج التأسيسي لقسم اللغة الإنجليزية في كلية التقنية [redacted]. مشاركتكم في هذه الدراسة وآراؤكم حول برنامج كيزلت (Quizlet) سيساعد الباحث في معرفة الطريقة المثلى لتدريس وتعلم مفردات اللغة الانجليزية.

هل يجب علينا المشاركة في الدراسة؟

مشاركة الطلاب والطالبات في الدراسة اختيارية، ولن يكون لها أي علاقة بتحصيل الدرجات الدراسية. كما أن المعلومات التي يتم الحصول عليها من المشاركين لن يطلع عليها المعلمين او المعلمات.
الطلاب والطالبات اللذين يتم اختيارهم للمشاركة في الدراسة سوف يتم إعطائهم هذا النموذج لتزويدهم بالمعلومات الخاصة بالبرنامج والتأكيد على أن آلية المشاركة اختيارية، ومن ثم التوقيع بالموافقة على المشاركة في البرنامج. وإذا تمت مشاركتهم في الدراسة فإن لهم الحق في الانسحاب من الدراسة في أي وقت وبدون إبداء أية أسباب للانسحاب فقط الاعتذار عن طريق إخبار الباحثة، منى مقيبل، بالاتصال على رقم الهاتف: او إرسال رسالة الى البريد الإلكتروني التالي: m.h.a.muqaibal@pgr.reading.ac.uk .

ما هي الإجراءات التي سيتم إتباعها خلال مشاركة الطلاب والطالبات في الدراسة؟

المشاركة في هذه الدراسة سوف تنطوي على تنفيذ نشاط تعلم المفردات الإنجليزية، وذلك باستخدام برنامج كيزلت (Quizlet). سيقوم الطلبة بهذا النشاط كجزء من العملية التعليمة وذلك من خلال حصصهم الاعتيادية، ولكن سوف يطلب منهم تسجيل موافقتهم على تبادل البيانات الخاصة بهم مع الباحث. بالإضافة الى ذلك سيتم دعوة المشاركين للإجابة على الاستبيان الخاص بالدراسة، سيستغرق الاستبيان 10 دقائق كحد أقصى لإنجازه مع توفير بيئة مريحة لذلك. ومن الجدير بالذكر انه سيتم استخدام المعلومات التي جمعها الباحث لتحليل البيانات لأغراض البحث العلمي.

ما هي الفائدة المرجوة من المشاركة في البرنامج؟

المشاركون في البرنامج سوف تتاح لهم الفرصة لطرح آرائهم حول برنامج كويزلت (Quizlet)، ومساعدة الباحثة في معرفة أفضل الطرق لتدريس وتعلم المفردات الإنجليزية.

هل المعلومات سوف تكون سرية، وماذا سيحدث لنتائج البحث؟

جميع البيانات التي يتم جمعها سوف تكون سرية، ولن تستخدم أسماء حقيقية في هذه الدراسة أو في أي منشورات لاحقة. كما أن جميع الاستبانات والتسجيلات الصوتية سوف يتم حفظها في مكان خاص.

تعتبر البيانات التي يتم جمعها في هذه الدراسة معلومات أساسية لدراستي لمرحلة الدكتوراه. وسيتم نشر رسالة الدكتوراه بغرض الفائدة العلمية. كما ستستخدم هذه الدراسة في أهداف علمية مثل إصدار مقالات وكتب ومحاضرات. وستبقى جميع المعلومات التي يتم جمعها محفوظة بسرية تامة.

من هي اللجنة التي قامت بمراجعة هذا النموذج وصحة المعلومات؟

لقد تمت مراجعة هذا النموذج وقراءته من قبل لجنة البحوث في جامعة ريدنج في المملكة المتحدة (المعلومات متوفرة عند الطلب).

ما هي الإجراءات المتبعة في حال انسحاب المشاركين في الدراسة أثناء مشاركتهم؟

يمكن للطلاب والطالبات خلال البحث تغيير رأيهم والانسحاب في أي وقت دون أي تبعات. وإذا حدث ذلك بعد جمع البيانات، فإنه سوف يتم إزالة البيانات الخاصة بهم من الدراسة.

ما هي الإجراءات المتبعة في حال وقوع أي مشكلة؟

في هذه الحالة لا داعي للقلق أو الشكوى، يمكنك التواصل مع مشرفة الباحثة: البروفيسور كاثرين تيسوت، جامعة ريدنج في المملكة المتحدة،

رقم الهاتف: 0118 378 2674، البريد الإلكتروني: c.tissot@reading.ac.uk

شكرا جزيلاً لتعاونك وقراءة النموذج.

<u>نموذج موافقة الطلاب والطالبات على المشاركة في الدراسة</u>

عنوان الدراسة: تقويم فاعلية تعليم اللغة بنظام التكرار المتباعد باستخدام الحاسوب في تدريس وتعلم المفردات الإنجليزية في الحصص الدراسية: دراسة حالة في سلطنة عمان.

لقد قرأت ورقة المعلومات الخاصة بالدراسة وحصلت على نسخة منها. وأنا أفهم ما هو الغرض من هذه الدراسة.

اسم الطالب/ـة: ....

1. أوافق على المشاركة في الدراسة ومشاركة بياناتي مع الباحث.	نعم ________	لا ________
2. أوافق على إعطاء الباحثة الفرصة لطرح الأسئلة علي من خلال الاستبانة.	نعم ________	لا ________
3. أوافق على إعطاء الباحثة الإذن باستخدام التسجيل الصوتي.	نعم ________	لا ________

التوقيع: ....

التاريخ: ....

اسم الباحثة: منى القبيل.
عنوان البريد الإلكتروني للباحثة:

Appendix 7 Questionnaire (English version)

Researcher: Muna Muqaibal

Students' Language Background and Attitude Towards Using 'Quizlet' Flashcards in Vocabulary Learning

The purpose of this questionnaire is to collect information about students' attitude towards using the software programme 'Quizlet' for English vocabulary learning. The questionnaire should take no more than 10 minutes. You are kindly requested to decide to what extent you agree with each statement using the following scale:

Strongly disagree = 1, Disagree = 2, Not sure = 3, Agree = 4 and Strongly agree = 5

For example:

No.	Statement	Strongly Disagree 1	2	3	4	Strongly Agree 5
1.	Summer holiday is too long.					

e.g. If you want a much longer summer holiday then you would tick (√) one for **strongly disagree**.

Date:

University of Reading

Questionnaire

PART A: Background Information
Mark in the appropriate space provided.

1. Student ID: _____

2. Age: _____

3. Gender: ☐ Male ☐ Female

4. School (precollege education): ☐ Public ☐ Private ☐ International

5. How many languages do you speak?.....

6. What is your native language? ☐ Arabic ☐ English ☐ Other local language: .....

7. What language(s) do you speak at home? ☐ Arabic ☐ English ☐ Other local language: .....

7a. If more than one language spoken at home, which language did you learn to speak first?
.....

8. In what language(s) did you receive the majority of your pre-college education?
☐ Arabic ☐ English

9. How much amount of time do you spend in **speaking English** at home?
☐ **0**=Never
☐ **1**=Rarely (1-30 % of the time)
☐ **2**=Some (30-60 % of the time)
☐ **3**=Most of the time (61- 90% of the time)
☐ **4**=All of the time

10. Have you ever been to an English-speaking region for the purpose of studying English?
☐ Yes ☐ No

10a. If yes, when?..... 10b. Where?.....

10c. For how long? ☐ 1semester or less ☐ 2 semesters ☐ more than 2 semesters

11. In the box below, rate your language ability in English language. Use the following rating:
1) Poor, 2) Fair, 3) Good, 4) Very good.

Listening	Speaking	Reading	Writing

12. What is your major?
☐ Business ☐ Engineering ☐ Information Communication Technology

PART B: Attitude towards Using Quizlet in Vocabulary Learning

Please read each statement and indicate the degree to which the statements apply to you using the following scale:

1 = Strongly Disagree 2 = Disagree 3 = Not sure 4 = Agree 5 = Strongly Agree

No.	Statement	Strongly Disagree 1	2	3	4	Strongly Agree 5
1.	I was able to learn English vocabulary more quickly with Quizlet than my own personal way of learning English vocabulary.					
2.	I was NOT satisfied with Quizlet as a good website to learn vocabulary.					
3.	I intend to study English vocabulary with Quizlet in the future.					
4.	Using Quizlet did NOT add anything to my English vocabulary knowledge.					
5.	I would not use Quizlet to study English vocabulary in the future.					
6.	It was easy for me to become skilful at using Quizlet.					
7.	If I am offered, I intend to study English vocabulary with Quizlet.					
8.	Using Quizlet for vocabulary learning is enjoyable.					
9.	It was difficult for me to study English vocabulary with Quizlet.					
10.	I would recommend Quizlet to learners.					
11	The L1 translations of target words in Quizlet were clear enough for me to understand new words.					
12.	The Quizlet website was clear and understandable.					
13.	I would recommend Quizlet to teachers.					
14.	Photos used in Quizlet were useful for me to remember new words					
15.	Using Quizlet is effortless.					
16.	I did NOT notice any inconsistencies, when I used Quizlet.					
17.	It is motivating to use Quizlet to learn new words.					
18.	Quizlet helps me to recover from mistakes quickly and easily.					
19.	I would NOT recommend Quizlet to anyone.					
20.	Using Quizlet improved my English vocabulary.					
21.	Both occasional and regular users would like Quizlet.					
22.	I was satisfied with Quizlet as a good website to learn vocabulary.					
23.	I think Quizlet was useful in my class.					
24.	Using Quizlet for vocabulary learning was boring.					
25.	It was easy for me to study English vocabulary with Quizlet.					

PART C: Open-ended Questions.

1. What aspect of Quizlet did you like the most? Why?

2. How easy did you find Quizlet to use?

3. If you have a choice would you prefer the college to use Quizlet for vocabulary?

Appendix 8 Questionnaire (Arabic version)

Researcher: Muna Muqaibal

الخلفية اللغوية للطلبة وآرائهم حول برنامج كيزلت (Quizlet) لتعلُّم المفردات الإنجليزية

الغرض من هذا الاستبيان هو جمع المعلومات عن **آراء** الطلبة حول استخدام برنامج "كيزلت" لتعلُّم المفردات الإنجليزية. حيث سيستغرق الاستبيان 10 دقائق كحد أقصى لإنجازه. يرجى من المشارك في الاستبيان تحديد إلى أي مدى يمكن ان يتفق مع كل جملة باستخدام المقياس التالي:

لا أوافق بشدة = 1 لا أوافق = 2 غير متأكد = 3 أوافق = 4 أوافق بشدة = 5

مثال:

رقم	الجملة	لا أوافق بشدة 1	2	3	4	أوافق بشدة 5
.1	إجازة الصيف طويلة جدًا.					

مثلا: إذا كنت ترغب في أن تكون عطلة الصيف أطول، فمن المرجح ان تضع علامة (√) أسفل الرقم واحد (**لا أوافق بشدة**).

التاريخ:.....

استبيان

الجزء الاول: البيانات الشخصية

. رقم الطالب الأكاديمي: _____

. العمر: _____

. النوع: ذكر أنثى

. المدرسة (ما قبل المرحلة الجامعية): مدرسة حكومية مدرسة خاصة مدرسة عالمية

. كم عدد اللغات التي تتحدثها؟.....

. ماهي لغتك الام؟ العربية الإنجليزية لغة محلية أخرى:.....

. ماهي اللغة (اللغات) التي تتحدث بها في المنزل؟ العربية الإنجليزية لغة محلية أخرى:.....

7أ) اذا كنت تتحدث أكثر من لغة في المنزل، فما هي اللغة التي تعلمتها أولاً؟ .....

. ما هي اللغة (اللغات) التي تلقيت فيها معظم تعليمك ما قبل الجامعي؟ العربية الإنجليزية

. ما نسبة تحدثك باللغة الإنجليزية في المنزل؟
0 = أبدا
1 = نادرا (1-30٪ من الوقت)
2 = بعض الاحيان (30-60٪ من الوقت)
3 = معظم الوقت (61-90٪ من الوقت)
4 = كل الوقت

. هل سبق لك دراسة اللغة الإنجليزية في احدى الدول الناطقة بالإنجليزية ؟
نعم لا

10 أ) إذا كانت الإجابة بنعم، فمتى درست؟.....، 10ب) وأين؟.....

10ج) مدة الدراسة؟ فصل دراسي واحد فصليين دراسيين أكثر من فصليين

. في الجدول أدناه، قيم قدرتك اللغوية في اللغة الإنجليزية باستخدام التقييم التالي:
) ضعيف 2) متوسط 3) جيد 4) جيد جدا

Writing (الكتابة)	**Reading** (القراءة)	**Speaking** (التحدث)	**Listening** (الاستماع)

الجزء الثاني: استطلاع أراء الطلبة تجاه استخدام برنامج كيزلت (Quizlet) في تعلم المفردات الإنجليزية.

يرجى منك تحديد إلى أي مدى تتفق مع كل جملة باستخدام المقياس التالي:

لا أوافق بشدة = 1 لا أوافق = 2 غير متأكد = 3 أوافق = 4 أوافق بشدة = 5

رقم	الجمل	لا أوافق بشدة 1	2	3	4	أوافق بشدة 5
.1	كنت قادراً على تعلم المفردات الإنجليزية بشكل سرع باستخدام كيزلت (Quizlet)، أسرع من طريقتي الشخصية في تعلم المفردات الإنجليزية.					
.2	لم أكن راضيا عن كيزلت (Quizlet)، كموقع جيد لتعلم المفردات.					
.3	أعتزم دراسة المفردات الإنجليزية مع كيزلت (Quizlet) في المستقبل.					
.4	استخدام كيزلت (Quizlet) لم يضف أي شيء إلى حصيلتي المعرفية في المفردات الإنجليزية.					
.5	أنا لن أستخدم كيزلت (Quizlet) لدراسة المفردات الإنجليزية في المستقبل.					
.6	كان من السهل بالنسبة لي أن أصبح ماهراً في استخدام كيزلت (Quizlet).					
.7	إذا توفر لي برنامج كيزلت، فأني أعتزم دراسة المفردات الإنجليزية باستخدام كيزلت (Quizlet).					
.8	استخدام كيزلت (Quizlet) لتعلم المفردات الإنجليزية ممتع.					
.9	كان من الصعب بالنسبة لي دراسة المفردات الإنجليزية باستخدام كيزلت (Quizlet).					
.10	أود أن أنصح الطلبة باستخدام كيزلت (Quizlet).					
.11	كانت ترجمة الكلمات إلى اللغة العربية واضحة لي بما فيه الكفاية لفهم الكلمات الجديدة.					
.12	كان موقع كيزلت (Quizlet) واضحا ومفهوما.					
.13	أود أن أوصي المعلمين باستخدام كيزلت (Quizlet).					
.14	كانت الصور مفيدة بالنسبة لي لتذكر الكلمات الجديدة.					
.15	استخدام كيزلت (Quizlet) لا يتطلب جهداً.					
.16	لم ألاحظ أي تعارض في معاني الكلمات، عند استخدام كيزلت (Quizlet).					
.17	استخدام كيزلت (Quizlet) محفز لتعلم كلمات جديدة.					
.18	يمكنني تدارك الأخطاء بسرعة وسهولة مع استخدام كيزلت (Quizlet).					
.19	انا لا أنصح أي شخص باستخدام كيزلت (Quizlet).					
.20	استخدام كيزلت (Quizlet) طور حصيلتي من المفردات الإنجليزية.					
.21	كل من المستخدمين المؤقتين والمنتظمين يرغبون باستخدام كيزلت (Quizlet).					
.22	كنت راضيا عن كيزلت (Quizlet) كموقع جيد لتعلم المفردات الإنجليزية.					
.23	أعتقد إن كيزلت (Quizlet) كان مفيداً في الفصل الدراسي.					
.24	استخدام كيزلت (Quizlet) لتعلم المفردات ممل.					
.25	كان من السهل بالنسبة لي دراسة المفردات الإنجليزية باستخدام كيزلت (Quizlet).					

الجزء الثالث: أسئلة مفتوحة.

1. ما هو أكثر جانب أعجبك في استخدام **Quizlet**؟ ولماذا؟

.....
.....
.....
.....

2. كيف وجدت استخدام **Quizlet** في الفصل من حيث سهولة الاستخدام؟

.....
.....
.....
.....

3. . ماذا لو كنت صاحب القرار أو الاختيار، هل ستفضل استخدم الكلية لبرنامج **Quizlet** لتعلم المفردات؟

.....
.....
.....
.....

Appendix 9 Interview Questions

Indicative Interview Questions for teachers

The following questions will be asked during the interview:

1. What methods or strategies do you use in the classroom that helps improve vocabulary knowledge? Is technology a part of any vocabulary teaching/instruction?

2. Is Quizlet useful for English vocabulary learning? If so, how?

3. Which aspects of the programme did you find most useful? Which aspects did you find least useful?

4. Do you think Quizlet motivates students to learn English vocabulary? If yes, how? If no, what are the reasons? Give specific reasons.

5. Do you think students improved their vocabulary knowledge with Quizlet? If yes, how? If no, what are the reasons? Give specific reasons.

6. How did you find using Quizlet in the classroom for both teachers and learners in terms of the ease of use?

    a. Installation, implementation, and during the class.

7. How did you find L1 translation and photos provided in the programme?

    a. Are they clear enough for students?

8. Would you prefer to use Quizlet for your future vocabulary teaching? If so, how?

9. Would you recommend using Quizlet to other teachers?

10. Would you recommend using Quizlet to students outside the classroom? If so, how?

Appendix 10 Intervention Tests (Versions 1, 2 and 3)

Version 1

Test 1: Translate the following words into English.

1. _____ مقيم
2. _____ مدير مدرسة
3. _____ محيط
4. _____ مثال
5. _____ صورة
6. _____ درجة الحرارة
7. _____ الكرة الارضية
8. _____ الطقس
9. _____ أسلوب

Test 2: Translate the following words into Arabic.

1. access _____
2. area _____
3. balance _____
4. community _____
5. education_____
6. environment_____
7. hunter _____
8. journey _____

Test 3: Select the correct English equivalent for each of the following words and circle it.

1. مدير
   a. manager b. writer c. connector d. agreement
2. مهارة
   a. skill b. sleep c. management d. article
3. حل
   a. solution b. playground c. taste d. ball
4. عاصفة
   a. rain b. goal c. cloud d. storm
5. فريق
   a. pupil b. team c. trade d. channel
6. أدوات
   a. tools b. vehicles c. plates d. flags
7. مشروع
   a. space b. project c. import d. perfume
8. مستهلك
   a. actor b. organizer c. consumer d. chapter
9. سياح
   a. directors b. tourists c. businessmen d. accountants

Test 4: Select the correct L1 translation for each of the following words and circle it.

1. device
   a. مطبخ b. اختيار c. مجلد d. جهاز
2. information
   a. معلومات b. رسائل c. اخبار d. تقارير
3. economy
   a. اقتصاد b. اللحم c.قلم d. تجارة
4. century
   a. ألماس b. قرن c. مدينة d. عقد
5. flood
   a. مطر b. ريح c. فيضان d. جفاف
6. benefit
   a. فائدة b.زنجبيل c. يد d. مصنع
7. aid
   a. طلاب b. مساعدة c. تصليح d. قلم
8. export
   a. تصدير b. استيراد c. ابحار d. سفن

Version 2

Test 1: Translate the following words into English.

1. _____ الطقس
2. _____ درجة الحرارة
3. _____ مدير مدرسة
4. _____ صورة
5. _____ الكرة الارضية
6. _____ أسلوب
7. _____ محيط
8. _____ مثال
9. _____ مقيم

Test 2: Translate the following words into Arabic.

1. journey _____
2. environment_____
3. balance _____
4. community _____
5. access _____
6. education_____
7. area _____
8. hunter _____

Test 3: Select the correct English equivalent for each of the following words and circle it.

1. أدوات
   a. vehicles b. tools c. flags d. plates
2. حل
   a. ball b. taste c. playground d. solution
3. سياح
   a. tourists b. directors c. accountants d. businessmen
4. عاصفة
   a. cloud b. storm c. rain d. goal
5. فريق
   a. team b. pupil c. channel d. trade
6. مشروع
   a. import b. perfume c. space d. project
7. مدير
   a. writer b. manager c. agreement d. connector
8. مهارة
   a. management b. article c. skill d. sleep
9. مستهلك
   a. chapter b. consumer c. actor d. organizer

Test 4: Select the correct L1 translation for each of the following words and circle it.

1. aid
   a. قلم b. تصليح c. مساعدة d. طلاب
2. benefit
   a. يد b. مصنع c. فائدة d. زنجبيل
3. century
   a. مدينة b. عقد c. ألماس d. قرن
4. device
   a. مجلد b. جهاز c. مطبخ d. اختيار
5. economy
   a. قلم b. تجارة c. اقتصاد d. اللحم
6. export
   a. ابحار b. سفن c. تصدير d. استيراد
7. flood
   a. ريح b. جفاف c. مطر d. فيضان
8. information
   a. رسائل b. معلومات c. تقارير d. اخبار

Version 3

Test 1: Translate the following words into English.

1. _____ أسلوب
2. _____ مقيم
3. _____ الطقس
4. _____ درجة الحرارة
5. _____ الكرة الأرضية
6. _____ مثال
7. _____ صورة
8. _____ مدير مدرسة
9. _____ محيط

Test 2: Translate the following words into Arabic.

1. community _____
2. education _____
3. hunter _____
4. journey _____
5. environment _____
6. area_____
7. balance_____
8. access_____

Test 3: Select the correct English equivalent for each of the following words and circle it.

1. مشروع
   a. project b. space c. perfume d. import
2. أدوات
   a. plates b. flags c. tools d. vehicles
3. مدير
   a. writer b. agreement c. manager d. connector
4. مهارة
   a. article b. management c. sleep d. skill
5. سياح
   a. accountants b. businessmen c. tourists d. directors
6. حل
   a. taste b. ball c. solution d. playground
7. فريق
   a. trade b. channel c. pupil d. team
8. عاصفة
   a. storm b. cloud c. goal d. rain
9. مستهلك
   a. organizer b. chapter c. consumer d. actor

Test 4: Select the correct translation for each of the following words and circle it.

1. flood
   a. فيضان b. مطر c. ريح d. جفاف
2. export
   a. سفن b.ابحار c. استيراد d. تصدير
3. device
   a. اختيار b. مطبخ c. جهاز d. مجلد
4. aid
   a. مساعدة b. طلاب c. قلم d. تصليح
5. information
   a. معلومات b.رسائل c.اخبار d. تقارير
6. benefit
   a. مصنع b. يد c. زنجبيل d. فائدة
7. economy
   a. تجارة b. قلم c. اللحم d. اقتصاد
8. century
   a.عقد b. مدينة c. قرن d. ألماس

Appendix 11 Working Memory Test and Research Protocol

Research Protocol
Digit span tests

Subtest 1: Number Memory Forward
(Rote learning & memory/Attention/Encoding/Auditory Processing)
This subtest is designed to show how well the student can retain simple sequences of auditory information.

Requires
- Each test component has 8 items with trials of same length.
- Make sure the student feel comfortable and secure, and there are not any interruptions or noise around.
- Read each trial verbatim at the rate of one digit per second

Materials
1- Administration and scoring manual
2- Record form
3- Voice recorder
4- Attendance sheet

Note: there are no time limits, but examiner must be mindful of the rate of passing seconds

Procedures
- On arrival ask the student for their name and group number to assign them in their correct group list (massed or spaced). Check that they have already signed the consent and ask the student to sign the test attendance sheet.
- Start portable voice recorder. Say the student's name, the group name (massed or spaced) and the date.
- The researcher gives instructions for the subtest:

"I'm going to say sets of numbers; when I'm finished with each set, you repeat them back to me in the same order as you heard them."
"Don't worry if you can't remember everything, but try to say as much as you can and to speak clearly."
"First we'll practice. Listen carefully; I can't repeat them once we start. Ready?"

Practice the task:
Make sure the student can hear you well and know what is needed. First, the student need to be trained on the subtest for ' one attempt'.
Example A.
Say: **" Three** (pause) **Eight** (pause). **What numbers did you hear?"**
When the student finished, say **"Ok, now we'll do the test".**

The subtest 1 (Forward) and scoring

الاختبار وحساب درجات الاختبار

Start with the example above. Discontinue when student has made two consecutive 0-point responds.

numbers الأرقام									The correct answer الإجابة الصحيحة	Score الدرجات
6	4									
2	5									
3	1	6								
7	4	9								
6	9	5	7							
3	6	2	9							
8	3	9	4	6						
5	1	7	2	9						
4	2	5	1	8	7					
5	8	4	9	3	6					
1	5	2	8	4	9	7				
8	2	4	7	3	6	1				
9	3	7	5	1	6	4	8			
2	6	4	8	3	7	1	5			
3	8	1	9	5	2	7	4	6		
6	9	5	3	8	1	4	7	2		

Score

1. Score 1 point if the student gives a correct response.
2. Score 0 points if the student gives an incorrect response, say that they do not know the answer, or does not respond within approximately 30 seconds.

Test Protocol
Digit span tests

Subtest 2: Number Memory Reversed (Backward)
(working memory/Transformation of information/Mental manipulation/Visuo-spatial imaging)
This subtest is designed to show how well the student can retain and manipulate simple sequences of auditory information.

Requires
- Each test component has 8 items with trials of same length.
- Make sure the student feel comfortable and secure, and there are not any interruptions or noise around.
- Put a break between the two subtests (forward & backward).
- Read each trial verbatim at the rate of one digit per second

Materials
5- Administration and scoring manual
6- Record form
7- Voice recorder
8- Attendance sheet

Note: there are no time limits, but examiner must be mindful of the rate of passing seconds

Procedures
- On arrival ask the student for their name and group number to assign them in their correct group list (massed or spaced). Ask the student to sign the test attendance sheet.
- Start portable voice recorder. Say the student's name, the group name (massed or spaced) and the date.
- The researcher gives instructions for the subtest:

"I'm going to say sets of numbers; when I'm finished with each set, you repeat them back to me in reversed order."
"So if I say '4 – 1', you say '1-4'."
"First we'll practice. Listen carefully; I can't repeat them once we start. Ready?"

Practice the task:
Make sure the student can hear you well and know what is needed. First, the student need to be trained on the subtest for ' two attempts'.
Example A. 9 - 2
Say: **" Nine** (pause) **Two** (pause). **Now tell me the numbers in reversed order."** (2 – 9)
When the student finished, say **"Ok, now we'll do another set",** and go to Example B.
Example B. 1 - 6
Say: **" One** (pause) **Six** (pause). **"Now tell me the numbers in reversed order."** (6 – 1)
When the student finished, say **"Ok, now we'll go on the test items".**

The subtest 2 (Backward) and scoring

الاختبار وحساب درجات الاختبار

Start with the example above. Discontinue when student has made two consecutive 0-point responds.

numbers الأرقام									The correct answer الإجابة الصحيحة	Score الدرجات
7	3								3 7	
4	9								9 4	
5	2	8							8 2 5	
6	9	2							2 9 6	
4	9	5	3						3 5 9 4	
7	1	6	8						8 6 1 7	
3	7	5	8	1					1 8 5 7 3	
2	9	4	6	3					3 6 4 9 2	
8	2	5	1	9	4				4 9 1 5 2 8	
1	7	4	8	5	9				9 5 8 4 7 1	
4	9	1	7	3	5	8			8 5 3 7 1 9 4	
6	2	9	1	4	7	3			3 7 4 1 9 2 6	
8	1	6	4	9	7	2	5		5 2 7 9 4 6 1 8	
9	3	7	5	1	4	2	8		8 2 4 1 5 7 3 9	
3	6	8	4	2	7	1	5	9	9 5 1 7 2 4 8 6 3	
4	8	2	1	9	5	7	3	6	6 3 7 5 9 1 2 8 4	

Score

1. Score 1 point if the student gives a correct response.
2. Score 0 points if the student gives an incorrect response, say that they do not know the answer, or does not respond within approximately 30 seconds.

Appendix 12 Target words for the main study

Target Word list
Level 2 (Elementary Level) Vocabulary List (Nouns)

No.	Target words	L1 translation
Academic Word List (AWL)		
1.	Access	دخول
2.	Aid	مساعدة
3.	Area	منطقة
4.	Benefit	فائدة
5.	Community	مجتمع
6.	Consumer	مستهلك
7.	Device	جهاز
8.	Economy	اقتصاد
9.	Environment	بيئة
10.	Export	تصدير
11.	Globe	الكرة الارضية
12.	Image	صورة
13.	Instance	مثال
14.	Principal	مدير مدرسة
15.	Project	مشروع
16.	Resident	مقيم
17.	Style	أسلوب
18.	Team	فريق
2000 Level of Frequency		
19.	Balance	توازن
20.	Century	قرن
21.	Education	التعليم
22.	Flood	فيضان
23.	Hunter	صياد
24.	Information	معلومات
25.	Journey	رحلة
26.	Manager	مدير
27.	Ocean	محيط
28.	Skill	مهارة
29.	Solution	حل
30.	Storm	عاصفة
31.	Temperature	درجة الحرارة
32.	Tools	أدوات
33.	Tourists	سياح
34.	Weather	الطقس

Appendix 13 Vocabulary Level Tests (VLTs)

Version 2 The 2,000 word level

1 copy
2 event _____ end or highest point
3 motor _____ this moves a car
4 pity _____ thing made to be like another
5 profit
6 tip

1 accident
2 debt _____ loud deep sound
3 fortune _____ something you must pay
4 pride _____ having a high opinion of yourself
5 roar
6 thread

1 coffee
2 disease _____ money for work
3 justice _____ a piece of clothing
4 skirt _____ using the law in the right way
5 stage
6 wage

1 clerk
2 frame _____ a drink
3 noise _____ office worker
4 respect _____ unwanted sound
5 theater
6 wine

1 dozen
2 empire _____ chance
3 gift _____ twelve
4 opportunity _____ money paid to the government
5 relief
6 tax

1 admire
2 complain _____ make wider or longer
3 fix _____ bring in for the first time
4 hire _____ have a high opinion of someone
5 introduce
6 stretch

1 arrange
2 develop _____ grow
3 lean _____ put in order
4 owe _____ like more than something else
5 prefer
6 seize

1 blame
2 elect _____ make
3 jump _____ choose by voting
4 manufacture _____ become like water
5 melt
6 threaten

1 ancient
2 curious _____ not easy
3 difficult _____ very old
4 entire _____ related to God
5 holy
6 social

1 bitter
2 independent _____ beautiful
3 lovely _____ small
4 merry _____ liked by many people
5 popular
6 slight

Version 2 The 3,000 word level

1 bull
2 champion _____ formal and serious manner
3 dignity _____ winner of a sporting event
4 hell _____ building where valuable
5 museum objects are shown
6 solution

1 blanket
2 contest _____ holiday
3 generation _____ good quality
4 merit _____ wool covering used on
5 plot beds
6 vacation

1 comment
2 gown _____ long formal dress
3 import _____ goods from a foreign
4 nerve country
5 pasture _____ part of the body which
6 tradition carries feeling

1 administration
2 angel _____ group of animals
3 frost _____ spirit who serves God
4 herd _____ managing business and
5 fort affairs
6 pond

1 atmosphere
2 counsel _____ advice
3 factor _____ a place covered with grass
4 hen _____ female chicken
5 lawn
6 muscle

1 abandon
2 dwell _____ live in a place
3 oblige _____ follow in order to catch
4 pursue _____ leave something
5 quote permanently
6 resolve

1 assemble
2 attach _____ look closely
3 peer _____ stop doing something
4 quit _____ cry out loudly in fear
5 scream
6 toss

1 drift
2 endure _____ suffer patiently
3 grasp _____ join wool threads together
4 knit _____ hold firmly with your hands
5 register
6 tumble

1 brilliant
2 distinct _____ thin
3 magic _____ steady
4 naked _____ without clothes
5 slender
6 stable

1 aware
2 blank _____ usual
3 desperate _____ best or most important
4 normal _____ knowing what is happening
5 striking
6 supreme

Version 2 The 5,000 word level

1 analysis
2 curb
3 gravel
4 mortgage
5 scar
6 zeal

_____ eagerness
_____ loan to buy a house
_____ small stones mixed with sand

1 cavalry
2 eve
3 ham
4 mound
5 steak
6 switch

_____ small hill
_____ day or night before a holiday
_____ soldiers who fight from horses

1 circus
2 jungle
3 nomination
4 sermon
5 stool
6 trumpet

_____ musical instrument
_____ seat without a back or arms
_____ speech given by a priest in a church

1 artillery
2 creed
3 hydrogen
4 maple
5 pork
6 streak

_____ a kind of tree
_____ system of belief
_____ large gun on wheels

1 chart
2 forge
3 mansion
4 outfit
5 sample
6 volunteer

_____ map
_____ large beautiful house
_____ place where metals are made and shaped

1 contemplate
2 extract
3 gamble
4 launch
5 provoke
6 revive

_____ think about deeply
_____ bring back to health
_____ make someone angry

1 demonstrate
2 embarrass
3 heave
4 obscure
5 relax
6 shatter

_____ have a rest
_____ break suddenly into small pieces
_____ make someone feel shy or nervous

1 correspond
2 embroider
3 lurk
4 penetrate
5 prescribe
6 resent

_____ exchange letters
_____ hide and wait for someone
_____ feel angry about something

1 decent
2 frail
3 harsh
4 incredible
5 municipal
6 specific

_____ weak
_____ concerning a city
_____ difficult to believe

1 adequate
2 internal
3 mature
4 profound
5 solitary
6 tragic

_____ enough
_____ fully grown
_____ alone away from other things

Appendix 14 Vocabulary Log

Level 1 VOCABULARY LOG

#	WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 5	WEEK 6	WEEK 7	WEEK 8	WEEK 9	WEEK 10
1	Human	Dangerous	Protect	Nervous	Public	difficult	Vacation	recommend	Invent	Complete
2	City	Incredible	Help	Seek	Rises	Site	Earn	indicator	Movement	Regard
3	Job	Confident	Close	Risky	Carry	Modern	Building	Tag	Graduate	Reason
4	Male	Character	Enough	Succeed	Crowded	Glide	Young	Become	Engineer	Aim
5	Female	Situation	Hope	Extreme	Identical	Unusual	Develop	Competition	Tough	Inspire
6	Nation	Adventure	Speed	Activity	Behave	Clue	Dream	Tool	Famous	Century
7	Common	Information	Alone	Satisfy	Describe	Manmade	Agree	Electricity	History	Capital
8	Together	Business	Exciting	Career	weight	Leader	Company	Lifestyle	Include	Internal
9	Few	Adventure	Goal	Swim	Ride	Death	Happen	Rules	Machine	Gold
10	Different	Discoveries	Favourite	Bike	Health	Hidden	Knowledge	Rough	Steps	Explore
11	Increase	Challenge	Alike	Hike	Idea	Roll	Plan	Traditions	Huge	Outgoing
12	Million	Experience	Native	Act	Because	Spread	Save	Realize	Goal	Athlete
13	Instantly	Foreign	Earth	Drive	Afford to	Amazing	Relax	Afraid	Professor	Problem
14	Probably	Important	Special	Cost	Advice	Contest	Chance	Nature	Calculate	Sick
15	Connect	Population	Typical	Success	Prize	View	Adapt	Professional	Report	Unknown
16	However	Commute	Trouble	Similar	Costume	Direction	Spectacular	Climb	Grow up	Skill
17	Popular	Comfortable	Lazy	Safe	Perform	Missing	Almost	Encourage	Achieve	Attraction
18	Social	Carefree	Media	Adopt	Spend	Habits	Destroy	Environment	Design	Flood
19	Constant	Traveller	Impact	Urban	Possible	Customer	Effect	Activity	Create	Share
20	Change	Passengers	Easy	Shy	Return	Helmet	Predict	Equipment	Practice	Smart

Level 2 VOCABULARY LOG

NO.	WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 5	WEEK 6
1	Team	Popular	Attend	Primary	Information	Collect
2	Globe	Tourists	Decide to	Residents	Migrate	Recycle
3	Connected	Favorite	Independent	Principal	Recent	Solution
4	Result	Medium	Motivated	Similar	Trace	Clean up
5	Produce	Weak	Decide	Skill	Discover	Throw away
6	Take care of	Pretty	Aid	Persuade	Descendants	Deal With
7	Normal	Contribute	Available	Feast	Area	Combine
8	Communicate	Extraordinary	Exchange	Wedding	Common	Create
9	Project	Environment	Available	Outdoors	Journey	Proud
10	To arrive	Unexpected	Record	Entertain	Valuable	Garbage
11	Travel	Measure	Laughter	Enjoy	Sail	Forecast
12	Opportunity	Balance	Joke	Relax	Trade	Predict
13	Experiences	Depend on	Funny	To inform	Goods	Storm
14	Skills	Realize	Situation	Education	Ship	Temperature
15	Dangerous	Adventure	Unique	To Improve	Image	Flood
16	Creative	Although	Benefits	Hobby	Objects	Destroy
17	Explore	Physical	Exercise	Vacation	Look like	Cloudy
18	Help	Manager	Healthy	Personality	Rule	Weather
19	Believe	Pollution	Drawback	Appearance	Near by	Coast
20	Effect	Ocean	To solve	Overweight	Tools	Melt

No.	WEEK 7	WEEK 8	WEEK 9	WEEK 10
1	Ordinary	Company	Century	Melt
2	Decide to	Material	Appreciate	Community
3	Independent	Despite	Honest	Profit
4	Motivated	Proud	Consider	Ecology
5	Attend	Organization	Consumer	Sustainable
6	Solve	Be aware of	Ambitious	Trust
7	Record	Damage	Economy	Corporate
8	Available	Especially	Entire	Economy
9	Exchange	Spicy	Suggest	Connect
10	Migrate	Mild	For instance	Access
11	Ancestor	Raw	Desert	Device
12	Discover	Taste	Harmful	Contact
13	Ancient	Delicious	Export	Represent
14	Alive	Allow	Import	Prevent
15	Remain	Nutritious	Local	Garbage
16	Probably	Popular	Comfortable	Realize
17	Hunter	Neighbourhood	Residents	Reduce
18	Reach	Numerous	Style	Response
19	Area	Serve	Spacious	Get rid of
20	Cause	Crunchy	Building	Sensible

Appendix 15 Participants' Background Information

Background Information		Control N (%)	Intensive N (%)	Spaced N (%)	Total
Gender	Male	26 (79)	25 (73.5)	23 (79)	**74 (77%)**
	Female	7 (21)	9 (26.5)	6 (21)	**22 (23%)**
Precollege Education	Public	23 (70)	28 (82)	22 (76)	**73 (76%)**
	Private	10 (30)	6 (18)	7 (24)	**23 (24%)**
	International	0	0	0	**0**
Speaking English at home	Never	13 (39)	27 (79)	13 (45)	**53 (55%)**
	Rarely	18 (55)	6 (18)	14 (48)	**38 (40%)**
	Sometimes	2 (6)	1 (3)	2 (7)	**5 (5%)**
	Most of the time	0	0	0	**0**
	All of the time	0	0	0	**0**
Language Background	Monolingual	23 (70)	8 (23.5)	20 (69)	**51 (53%)**
	Bilingual	10 (30)	25 (73.5)	9 (31)	**44 (46%)**

Appendix 16 Non-parametric Test (Kruskal-Wallis test)

The 4-day sub-groups

Kruskal-Wallis test of Recognition and Recall between 4-day sub-groups (Control, Intensive and Spaced) at each time point.

Parameter		*H*	*df*	*p*
Recognition	Pre-test	.922	2	.162
	Immediate post-test	30.742	2	.001
	Delayed post-test	26.491	2	.001
Recall	Pre-test	3.170	2	.205
	Immediate post-test	28.592	2	.001
	Delayed post-test	28.807	2	.001

Pairwise comparisons of Recognition and Recall between 4-day sub-groups (Control, Intensive and Spaced**)** at **Immediate post-test**, with adjusted *p*-values.

	Groups	*z*	$\sqrt{N}$	*r*	*p*. value
Recognition	**Control-Spaced**	-4.58	5.48	-.84	.001
	Control-Intensive	-4.97	5.74	-.87	.001
	Spaced-Intensive	0.15	5.57	.03	1.000
Recall	**Control-Spaced**	-4.23	5.48	-.77	.001
	Control-Intensive	-4.92	5.74	-.86	.001
	Spaced-Intensive	0.46	5.57	.08	1.000

Pairwise comparisons of Recognition and Recall between 4-day sub-groups (Control, Intensive and Spaced) at **Delayed post-test**, with adjusted *p*-values.

	Groups	*z*	$\sqrt{N}$	*r*	*p*. value
Recognition	**Control-Spaced**	-4.27	5.48	-.78	.001
	Control-Intensive	-4.60	5.74	-.80	.001
	Spaced-Intensive	0.11	5.57	.02	1.000
Recall	**Control-Spaced**	-4.45	5.48	-.81	.001
	Control-Intensive	-4.80	5.74	-.84	.001
	Spaced-Intensive	0.12	5.57	.02	1.000

The 28-day sub-groups

Kruskal-Wallis test of Recognition and Recall between 28-day sub-groups (Control, Intensive and Spaced) at each time point.

Parameter		*H*	*df*	*p*
Recognition	Pre-test	1.297	2	.523
	Immediate post-test	29.465	2	.001
	Delayed post-test	26.621	2	.001
Recall	Pre-test	.497	2	.780
	Immediate post-test	19.374	2	.001
	Delayed post-test	19.212	2	.001

Pairwise comparisons of Recognition and Recall between 28-day sub-groups (Control, Intensive and Spaced) at **Immediate post-test**, with adjusted *p*-values.

	Groups	*z*	$\sqrt{N}$	*r*	*p*. value
Recognition	**Control-Spaced**	-4.41	5.66	-.78	.001
	Control-Intensive	-4.90	5.83	-.84	.001
	Spaced-Intensive	0.33	5.66	.06	1.000
Recall	**Control-Spaced**	-3.69	5.66	-.65	.001
	Control-Intensive	-3.88	5.83	-.67	.001
	Spaced-Intensive	-0.31	5.66	0.05	1.000

Pairwise comparisons of Recognition and Recall between 28-day sub-groups (Control, Intensive and Spaced) at **Delayed post-test**, with adjusted *p*-values.

	Groups	*z*	$\sqrt{N}$	*r*	*p*. value
Recognition	**Control-Spaced**	-4.33	5.66	-.77	.001
	Control-Intensive	-4.55	5.83	-.78	.001
	Spaced-Intensive	0.08	5.66	0.01	1.000
Recall	**Control-Spaced**	-3.63	5.66	-.64	.001
	Control-Intensive	-3.91	5.83	-.67	.001
	Spaced-Intensive	-0.39	5.66	-0.07	1.000

Appendix 17 Non-parametric Test (Friedman's ANOVA)

The 4-day sub-groups

Friedman's Two-Way Analysis of Recognition and Recall for **Control Group (**4-day RI sub-group) at each time point.

			N	x^2	*df*	*P.* value
Control	**Recognition**	Pre-test Immediate post-test Delayed post-test	16	6.60	2	.037
	Recall	Pre-test Immediate post-test Delayed post-test	16	1.37	2	.504

Friedman's Two-Way Analysis of Recognition and Recall for **Intensive Group (**4-day RI sub-group) at each time point.

			N	x^2	*df*	*P.* value
Intensive	**Recognition**	Pre-test Immediate post-test Delayed post-test	17	28.88	2	.001
	Recall	Pre-test Immediate post-test Delayed post-test	17	30.70	2	.001

Friedman's Two-Way Analysis of Recognition and Recall for **Spaced Group (**4-day RI sub-group) at each time point.

			N	x^2	***df***	***P*. value**
Spaced	**Recognition**	Pre-test Immediate post-test Delayed post-test	14	24.67	2	.001
	Recall	Pre-test Immediate post-test Delayed post-test	14	22.62	2	.001

The 28-day sub-groups

Friedman's Two-Way Analysis of Recognition and Recall for **Control Group (**28-day RI sub-group) at each time point.

			N	x^2	*df*	*P.* value
Control	**Recognition**	Pre-test Immediate post-test Delayed post-test	17	3.09	2	.214
	Recall	Pre-test Immediate post-test Delayed post-test	17	2.70	2	.260

Friedman's Two-Way Analysis of Recognition and Recall for **Intensive Group (**28-day RI sub-group) at each time point.

			N	x^2	*df*	*P.* value
Intensive	**Recognition**	Pre-test Immediate post-test Delayed post-test	17	28.77	2	.001
	Recall	Pre-test Immediate post-test Delayed post-test	17	26.59	2	.001

Friedman's Two-Way Analysis of Recognition and Recall for **Spaced Group (**28-day RI sub-group) at each time point.

			N	x^2	***df***	***P*. value**
Spaced	**Recognition**	Pre-test Immediate post-test Delayed post-test	15	25.47	2	.001
	Recall	Pre-test Immediate post-test Delayed post-test	15	28.53	2	.001

Appendix 18 Non-parametric Test (Kruskal-Wallis test of Passive/Active Recognition and Passive/Active Recall)

The 4-day sub-groups

Kruskal-Wallis test of Passive/Active Recognition and Passive/Active Recall between 4-day sub-groups (Control, Intensive and Spaced) at each time point.

Parameter		***H***	***df***	***p***
Passive Recognition	Pre-test	.113	2	.945
	Immediate post-test	33.401	2	.001
	Delayed post-test	28.089	2	.001
Active Recognition	Pre-test	.224	2	.894
	Immediate post-test	29.365	2	.001
	Delayed post-test	22.467	2	.001
Passive Recall	Pre-test	2.427	2	.297
	Immediate post-test	29.624	2	.001
	Delayed post-test	31.225	2	.001
Active Recall	Pre-test	2.637	2	.268
	Immediate post-test	24.206	2	.001
	Delayed post-test	23.435	2	.001

Pairwise comparisons of Passive/Active Recognition and Passive/Active Recall between 4-day sub-groups (Control, Intensive and Spaced**)** at **immediate post-test**, with adjusted *p*-values.

	Groups	***z***	$\sqrt{N}$	***r***	***p*. value**
Passive Recognition	Control-Intensive	-4.99	5.74	-.87	.001
	Control- Spaced	-4.99	5.48	-.91	.001
	Spaced-Intensive	0.24	5.57	.04	.808
Active Recognition	Control-Intensive	-4.47	5.74	-.78	.001
	Control- Spaced	-4.86	5.48	-.89	.001
	Spaced-Intensive	0.15	5.57	.03	1.000
Passive Recall	Control-Intensive	-4.47	5.74	-.78	.001
	Control- Spaced	-4.90	5.48	-.89	.001
	Spaced-Intensive	0.19	5.57	.03	1.000
Active Recall	Control-Intensive	-3.58	5.74	-.62	.001
	Control- Spaced	-4.70	5.48	-.86	.001
	Spaced-Intensive	0.91	5.57	.08	1.000

Pairwise comparisons of Passive/Active Recognition and Passive/Active Recall between 4-day sub-groups (Control, Intensive and Spaced) at **Delayed post-test**, with adjusted *p*-values.

	Groups	***z***	$\sqrt{N}$	***r***	***p*. value**
Passive Recognition	Control-Intensive	-4.62	5.74	-.80	.001
	Control- Spaced	-4.53	5.48	-.83	.001
	Spaced-Intensive	-0.13	5.57	-.02	1.000
Active Recognition	Control-Intensive	-4.27	5.74	-.74	.001
	Control- Spaced	-3.89	5.48	-.71	.001
	Spaced-Intensive	0.17	5.57	.03	1.000
Passive Recall	Control-Intensive	-4.77	5.74	-.83	.001
	Control- Spaced	-4.87	5.48	-.89	.001
	Spaced-Intensive	-0.34	5.57	-.06	1.000
Active Recall	Control-Intensive	-4.39	5.74	-.76	.001
	Control- Spaced	-3.93	5.48	-.72	.001
	Spaced-Intensive	0.25	5.57	.04	1.000

The 28-day sub-groups

Kruskal-Wallis test of Passive/Active Recognition and Passive/Active Recall between 28-day sub-groups (Control, Intensive and Spaced) at each time point.

Parameter		*H*	*df*	*p*
Passive Recognition	Pre-test	3.167	2	.205
	Immediate post-test	32.240	2	.001
	Delayed post-test	28.890	2	.001
Active Recognition	Pre-test	.062	2	.969
	Immediate post-test	21.006	2	.001
	Delayed post-test	20.044	2	.001
Passive Recall	Pre-test	.313	2	.855
	Immediate post-test	23.968	2	.001
	Delayed post-test	14.497	2	.001
Active Recall	Pre-test	.105	2	.949
	Immediate post-test	11.964	2	.003
	Delayed post-test	17.719	2	.001

Pairwise comparisons of Passive/Active Recognition and Passive/Active Recall between 28-day sub-groups (Control, Intensive and Spaced) at **immediate post-test**, with adjusted *p*-values.

	Groups	***z***	$\sqrt{N}$	***r***	***p*. value**
Passive Recognition	Control-Intensive	-5.16	5.83	-.89	.001
	Control- Spaced	-4.58	5.66	-.81	.001
	Spaced-Intensive	0.42	5.66	.07	1.000
Active Recognition	Control-Intensive	-4.18	5.83	-.72	.001
	Control- Spaced	-3.66	5.66	-.65	.001
	Spaced-Intensive	0.39	5.66	.07	1.000
Passive Recall	Control-Intensive	-4.13	5.83	-.71	.001
	Control- Spaced	-4.30	5.66	-.76	.001
	Spaced-Intensive	-0.31	5.66	-0.05	1.000
Active Recall	Control-Intensive	-2.80	5.83	-.48	.015
	Control- Spaced	-3.13	5.66	-.55	.005
	Spaced-Intensive	-0.42	5.66	-0.07	1.000

Pairwise comparisons of Passive/Active Recognition and Passive/Active Recall between 28-day sub-groups (Control, Intensive and Spaced) at **Delayed post-test**, with adjusted *p*-values.

	Groups	***z***	$\sqrt{N}$	***r***	***p*. value**
Passive Recognition	Control-Intensive	-4.61	5.83	-.79	.001
	Control- Spaced	-4.65	5.66	-.82	.001
	Spaced-Intensive	-0.19	5.66	-.03	1.000
Active Recognition	Control-Intensive	-4.02	5.83	-.69	.001
	Control- Spaced	-3.67	5.66	-.65	.001
	Spaced-Intensive	0.22	5.66	0.04	1.000
Passive Recall	Control-Intensive	-2.99	5.83	-.51	.001
	Control- Spaced	-3.51	5.66	-.62	.001
	Spaced-Intensive	-0.61	5.66	-0.11	1.000
Active Recall	Control-Intensive	-3.61	5.83	-.62	.001
	Control- Spaced	-3.65	5.66	-.64	.001
	Spaced-Intensive	-0.16	5.66	-0.03	1.000

Appendix 19 Non-parametric Test (Friedman's ANOVA of Passive/Active Recognition and Passive/Active Recall)

The 4-day sub-groups

Friedman's Two-Way Analysis of Passive/Active Recognition and Passive/Active Recall for Control Group (4-day RI sub-group) at each time point.

			N	x^2	*df*	*P.* value
Control	**Passive Recognition**	Pre-test Immediate post-test Delayed post-test	16	3.13	2	.210
	Active Recognition	Pre-test Immediate post-test Delayed post-test	16	10.41	2	.005
	Passive Recall	Pre-test Immediate post-test Delayed post-test	16	2.67	2	.264
	Active Recall	Pre-test Immediate post-test Delayed post-test	16	6.13	2	.047

Friedman's Two-Way Analysis of Passive/Active Recognition and Passive/Active Recall for **Intensive Group (**4-day RI sub-group) at each time point.

			N	x^2	*df*	*P*. value
Intensive	**Passive Recognition**	Pre-test Immediate post-test Delayed post-test	17	25.82	2	.001
	Active Recognition	Pre-test Immediate post-test Delayed post-test	17	30.47	2	.001
	Passive Recall	Pre-test Immediate post-test Delayed post-test	17	29.55	2	.001
	Active Recall	Pre-test Immediate post-test Delayed post-test	17	27.84	2	.001

Friedman's Two-Way Analysis of Passive/Active Recognition and Passive/Active Recall for **Spaced Group (**4-day RI sub-group) at each time point.

			N	x^2	*df*	*P.* value
Spaced	**Passive Recognition**	Pre-test Immediate post-test Delayed post-test	14	26.18	2	.001
	Active Recognition	Pre-test Immediate post-test Delayed post-test	14	23.09	2	.001
	Passive Recall	Pre-test Immediate post-test Delayed post-test	14	24.50	2	.001
	Active Recall	Pre-test Immediate post-test Delayed post-test	14	21.81	2	.001

The 28-day sub-groups

Friedman's Two-Way Analysis of Passive/Active Recognition and Passive/Active Recall for **Control Group (**28-day RI sub-group) at each time point.

			N	x^2	***df***	***P*. value**
Control	**Passive Recognition**	Pre-test Immediate post-test Delayed post-test	17	2.47	2	.291
	Active Recognition	Pre-test Immediate post-test Delayed post-test	17	3.25	2	.197
	Passive Recall	Pre-test Immediate post-test Delayed post-test	17	2.92	2	.232
	Active Recall	Pre-test Immediate post-test Delayed post-test	17	4.33	2	.115

Friedman's Two-Way Analysis of Passive/Active Recognition and Passive/Active Recall for **Intensive Group (**28-day RI sub-group) at each time point.

			N	x^2	*df*	*P*. value
Intensive	**Passive Recognition**	Pre-test Immediate post-test Delayed post-test	17	30.61	2	.001
	Active Recognition	Pre-test Immediate post-test Delayed post-test	17	30.56	2	.001
	Passive Recall	Pre-test Immediate post-test Delayed post-test	17	26.14	2	.001
	Active Recall	Pre-test Immediate post-test Delayed post-test	17	18.76	2	.001

Friedman's Two-Way Analysis of Passive/Active Recognition and Passive/Active Recall for **Spaced Group (**28-day RI sub-group) at each time point.

			N	x^2	*df*	*P*. value
Spaced	**Passive Recognition**	Pre-test Immediate post-test Delayed post-test	15	24.75	2	.001
	Active Recognition	Pre-test Immediate post-test Delayed post-test	15	27.04	2	.001
	Passive Recall	Pre-test Immediate post-test Delayed post-test	15	24.53	2	.001
	Active Recall	Pre-test Immediate post-test Delayed post-test	15	25.48	2	.001

Appendix 20 Logistic Plan for the Main Study

Logistic Plan

Week#1 23-27/9/2018	Class-1 (IG-1)[1]	Class-2 (IG-2)[2]	Class-3 (SG-1)[3]	Class-4 (SG-2)[4]	Class-5 (CG-1)[5]	Class-6 (CG-2)[6]	Comments/Obstacle/Actions
Sunday	- Students' consents (30mins)	- Students' consents (30mins)	- Students' consents (30mins)	- Students' consents (30mins)	- Students' consents (30mins)	- Students' consents (30mins)	
Monday	- Background Questionnaire (10mins)	- Background Questionnaire (10mins)	- Background Questionnaire (10mins)	- Background Questionnaire (10mins)	- Background Questionnaire (10mins)	- Background Questionnaire (10mins)	
Tuesday	- VLT (45mins)[7]	- VLT (45mins)	- VLT (45mins)	- VLT (45mins)	- VLT (45mins)	- VLT (45mins)	
Wednesday	- Introduce Quizlet (30mins)	- Introduce Quizlet (30mins)	- Introduce Quizlet (30mins)	- Introduce Quizlet (30mins)			
Thursday	- Register & log in (30mins)	- Register & log in (30mins)	- Register & log in (30mins)	- Register & log in (30mins)			
Week#2 30/9-4/10/2018	**Class-1 (IG-1)**	**Class-2 (IG-2)**	**Class-3 (SG-1)**	**Class-4 (SG-2)**	**Class-5 (CG-1)**	**Class-6 (CG-2)**	
Sunday	- Pre-test (34mins) - Intervention1 (20mins)	- Pre-test (34mins) - Intervention1 (20mins)	- Pre-test (34mins) - Intervention1 (20mins)	- Pre-test (34mins) - Intervention1 (20mins)	- Pre-test (34mins)	- Pre-test (34mins)	
Monday	- Intervention2 (20mins)	- Intervention2 (20mins)					
Tuesday	- Intervention3 (20mins)	- Intervention3 (20mins)					
Wednesday	- Intervention4 (20mins) -Immediate post-test (34mins)	- Intervention4 (20mins) -Immediate post-test (34mins)					
Thursday	- Questionnaire (10mins)	- Questionnaire (10mins)					

[1] IG-1=Intensive Group-Class 1
[2] IG-2=Intensive Group-Class 2
[3] SG-1=Spaced Group-Class 3
[4] SG-2=Spaced Group-Class 4
[5] CG-1=Control Group-Class 5
[6] CG-2=Control Group-Class 6
[7] VLT=Vocabulary Levels Tests

Week#3 7-11/10/2018	Class-1 (IG-1)	Class-2 (IG-2)	Class-3 (SG-1)	Class-4 (SG-2)	Class-5 (CG-1)	Class-6 (CG-2)	Comments/Obstacle/Actions
Sunday	- 4-day RI[1] Delayed post-test (34mins)-for All students		- Intervention2 (20mins)	- Intervention2 (20mins)	- 4-day RI Delayed post-test (34mins)-for ALL students		
Monday							
Tuesday							
Wednesday							
Thursday							

Week#4 14-18/10/2018	Class-1 (IG-1)	Class-2 (IG-2)	Class-3 (SG-1)	Class-4 (SG-2)	Class-5 (CG-1)	Class-6 (CG-2)	
Sunday			- Intervention3 (20mins)	- Intervention3 (20mins)			
Monday							
Tuesday							
Wednesday							
Thursday							

[1] RI=the Retention Interval

Week#5 21-25/10/2018	Class-1 (IG-1)	Class-2 (IG-2)	Class-3 (SG-1)	Class-4 (SG-2)	Class-5 (CG-1)	Class-6 (CG-2)	Comments/Obstacle/Actions
Sunday			- Intervention4 (20mins) -Immediate post-test (34mins)	- Intervention4 (20mins) -Immediate post-test (34mins)			
Monday			- Questionnaire (10mins)	- Questionnaire (10mins)			
Tuesday	Interview (30mins)		Interview (30mins)				
Wednesday		Interview (30mins)		Interview (30mins)			
Thursday			- 4-day RI Delayed post-test (34mins)-for ALL students			- 4-day RI Delayed post-test (34mins)-for ALL students	
Week#6 28/10-1/11/2018	Class-1 (IG-1)	Class-2 (IG-2)	Class-3 (SG-1)	Class-4 (SG-2)	Class-5 (CG-1)	Class-6 (CG-2)	
Sunday							
Monday							
Tuesday							
Wednesday		- 28-day RI Delayed post-test (34mins)-for ALL students			- 28-day RI Delayed post-test (34mins)-for ALL students		
Thursday							

Week#7 4-8/11/2018	Class-1 (IG-1)	Class-2 (IG-2)	Class-3 (SG-1)	Class-4 (SG-2)	Class-5 (CG-1)	Class-6 (CG-2)	Comments/Obstacle/Actions
Sunday							
Monday	- WM tests[1] (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	
Tuesday	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	
Wednesday	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	
Thursday	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	- WM tests (5 students per hour)	

Week#9 18/11/2018	Class-1 (IG-1)	Class-2 (IG-2)	Class-3 (SG-1)	Class-4 (SG-2)	Class-5 (CG-1)	Class-6 (CG-2)	Comments/Obstacle/Actions
Sunday				- 28-day RI Delayed post-test (34mins)-for ALL students		- 28-day RI Delayed post-test (34mins)-for ALL students	

[1] WM tests=Working memory tests. Students take working memory tests during their break between session. They usually have one hour break during the day.

Appendix 21 Target Words (Pilot study)

Target Word list
(2000-word level of frequency)

No.	Target words	L1 translation	2 syllables	Part of speech
1.	Approve	يوافق	Ap-prove	Verb
2.	Replace	يبدل	Re-place	Verb
3.	Attract	يجذب	At-tract	Verb
4.	Collect	يجمع	Col-lect	Verb
5.	Compare	يقارن	Com-pare	Verb
6.	Connect	يربط	Con-nect	Verb
7.	Explore	يكتشف	Ex-plore	Verb
8.	Improve	تحسن	Im-prove	Verb
9.	Display	يعرض	Dis-play	Verb
10.	Destroy	يدمر	Des-troy	Verb
11.	Balance	توازن	Bal-ance	Noun
12.	Message	رسالة	Mes-sage	Noun
13.	Diamond	ألماس	Dia-mond	Noun
14.	Chapter	فصل	Chap-ter	Noun
15.	Channel	قناة	Chan-nel	Noun
16.	Kitchen	مطبخ	Kitch-en	Noun
17.	Mistake	خطأ	Mis-take	Noun
18.	Sentence	جملة	Sen-tence	Noun
19.	Surface	سطح	Sur-face	Noun
20.	Weather	طقس	Weath-er	Noun

Appendix 22 Logistic Plan for the Pilot Study

Logistic Plan (Pilot Study)

	8:00 a.m.-9:00 a.m.	9:00-10:00	10:00-11:00
Sunday	Arriving in Oman early morning		
Monday 12/02/2018	Grp5[1]-ELC11 - Students' consents (20mins) - Pre-test (30mins) - Background Questionnaire (10mins)	Grp5-ELC11 VLT (50mins)	
Tuesday 13/02/2018	Grp7[2]-ELC13 - Students' consents (20mins) - Pre-test (30mins) - Background Questionnaire (10mins)	Grp7-ELC13 VLT (50mins)	Grp5-LAB2 Intervention1 (20mins)
Wednesday 14/02/2018			WM[3] tests (5 students-G5)
Thursday 15/02/2018			WM tests (5 students-G7)
Sunday 18/02/2018			WM tests (5 students-G5)
Monday 19/02/2018			WM tests (5 students-G7)
Tuesday 20/02/2018		Grp7-LAB3 - Intervention1&2 (40mins) - Immediate post-test (30mins)[4]	Grp5-LAB2 - Intervention2 (20mins) - Immediate post-test (30mins)
Wednesday 21/02/2018			WM tests (5 students-G5)
Thursday 22/02/2018			WM tests (5 students-G7)
Sunday 25/02/2018	Check the interview questions		WM tests
Monday 26/02/2018			WM tests
Tuesday 27/02/2018		Grp7-LAB3 - Delayed post-test (30mins) - Attitude Questionnaire (10mins)	Grp5-LAB2 - Delayed post-test (30mins) - Attitude Questionnaire (10mins)
Wednesday 28/02/2018			WM tests
Thursday			WM tests

[1] Massed Group
[2] Spaced Group
[3] WM tests =Work Memory tests (Digit Span)
[4] Someone else will be the invigilator as the researcher needs to conduct the 2nd intervention to Grp5 in LAB2. LAB3 is free from 10a.m.-11a.m on Tuesday.

Timetables

Time & Sessions for the Pilot Study

Week 1

HOUR	DAYS / TIMING	SUNDAY 11/02/2018	MONDAY 12/02/2018	TUESDAY 13/02/2018	WEDNESDAY 14/02/2018	THURSDAY 15/02/2018
1	*08:00-08:50*		*W* *MOHAMMED*	*W* *SHAROOQ*		
2	*09:00-09:50*		*W* *MOHAMMED*	*R* *LAB3* *SHAROOQ*		
3	*10:00-10:50*			*R* *MOHAMMED* *LAB2*		

Week 2

HOUR	DAYS / TIMING	SUNDAY 18/02/2018	MONDAY 19/02/2018	TUESDAY 20/02/2018	WEDNESDAY 21/02/2018	THURSDAY 22/02/2018
1	*08:00-08:50*					
2	*09:00-09:50*			*R* *LAB3* *SHAROOQ*		
3	*10:00-10:50*			*R* *MOHAMMED* *LAB2*		

Week3

HOUR	DAYS / TIMING	SUNDAY 25/02/2018	MONDAY 26/02/2018	TUESDAY 27/02/2018	WEDNESDAY 28/02/2018	THURSDAY 29/02/2018
1	*08:00-08:50*					
2	*09:00-09:50*			*R* *LAB3* *SHAROOQ (70mins)**		
3	*10:00-10:50*			*R* *MOHAMMED* *LAB2*		

Free periods for Grp5 & Grp7 students

Group : FDL 1-G 5

Tutor: MOHAMMED AL RAIMI ROOM NO: ELC-11 Massed Group

HOUR	DAYS / TIMING	SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY
1	*08:00-08:50*	*R* MOHAMMED	*W* MOHAMMED	*W* MOHAMMED	*W* MOHAMMED	*R* MOHAMMED
2	*09:00-09:50*	*R* MOHAMMED	*W* MOHAMMED	*SP* STIEVE	*W* MOHAMMED	*R* MOHAMMED
3	*10:00-10:50*			*R* MOHAMMED *LAB2*		*L* STIEVE
4	*11.00-11:50*	*L* STIEVE	*SP* LAB2 STIEVE		*L* STIEVE	
5	*12:00-12:50*	*L* STIEVE	*L* STIEVE		*SP* STIEVE	

ENGLISH LANGUAGE CENTER
STUDENT TIMETABLE

Group : FDL 1-G 7

Tutor: SHAROOQ ROOM NO: ELC-13 Spaced Group

HOUR	DAYS / TIMING	SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY
1	*08:00-08:50*	*R* SHAROOQ	*W* SHAROOQ	*W* SHAROOQ	*R* SHAROOQ	*W* SHAROOQ
2	*09:00-09:50*	*R* SHAROOQ	*W* SHAROOQ	*R* *LAB3* SHAROOQ	*W* SHAROOQ	*R* SHAROOQ
3	*10:00-10:50*			*L* CALIRE	*SP* HASHIM	
4	*11.00-11:50*	*L* CALIRE	*L* CALIRE	*SP* HASHIM		*SP* HASHAIM *LAB1*
5	*12:00-12:50*	*L* CALIRE	*L* CALIRE			

Appendix 23 Lesson Plans (Pilot Study)

Intervention Lesson Plan (Pilot Study)

Spaced Group

1st session

Date:	**Time:** 10:00 a.m.- 10:25 a.m.	**Warm-up**	**Task/Mode**	**Learning Method**	**Objectives**
Tuesday 13/02/2018	10 mins	Study flashcards Automatic-play for (3 mins) (Individual) Audio: On	Study flashcards Audio: On	- In pairs - Students can click to flip over the card, or use their arrow keys, and see the definition for that term. - Teacher guides students to click 'STAR' when they feel that the word is difficult. - After complete studying all words, students will be asked to study words with stars.	- Students learn the forms and their definitions (meanings).
	5 mins		Match	- Individual - Students will match all the target words with their meanings.	- Students match words with definitions to assess and build understanding of conceptual relationships. - Matching task requires students to evaluate, compare and match based on explicit, topic-specific relationships.
	10 mins		Spell (in English) Audio: On & Slow	- Individual - Students will listen to the word in English provided with its Arabic translation and image. - They will be asked to write the English equivalent.	- Students learn about phonological awareness and spelling.
Materials: PCs, headsets, overhead projector			**Avenue:** LAB 2		
Follow-up: 5 students meet the teacher individually in Room ELC21 for the WM test during the break.			**Assessment:**		

Spaced Group
2nd session

Date:	**Time:** 10:00 a.m.-10:25 a.m.	**Warm-up**	**Task/Mode**	**Learning Method**	**Objectives**
Tuesday 20/02/2018	10 mins	Study flashcards Automatic-play for (3 mins) Audio: On	Writing (writing the word forms in Arabic) Audio: On	- Individual - Teacher guides students to choose Writing mode on Quizlet and select Arabic. - When students write a word, they will automatically listen to the word in Arabic & English.	- Students learn the forms (oral & written) and write their definitions (meanings).
	5 mins		Writing (writing the word forms in English) Audio: On	- Individual - Students will match all the target words with their meanings.	- Students learn the forms (oral & written) and write their definitions (meanings).
	10 mins		Test	- Individual - Students will answer 5 written questions, 5 matching questions, 5 multiple choice questions & 5 True/False questions.	- Students revise words using different tasks.
Materials: PCs, headsets, overhead projector				**Avenue:** LAB 2	
Follow-up: Delayed post-test & Attitude Questionnaire one week later				**Assessment:** Immediate post-test (30 mins)	

Massed Group
1st & 2nd sessions

Date:	Time: 09:00 a.m.-10:20 a.m.	Warm-up	Task/Mode	Learning Method	Objectives
Tuesday 20/02/2018	10 mins	Study flashcards Automatic-play for (3 mins) Audio: On	Study flashcards Audio: On	- In pairs - Students can click to flip over the card, or use their arrow keys, and see the definition for that term. - Teacher guides students to click 'STAR' when they feel that the word is difficult. - After complete studying all words, students will be asked to study words with stars.	- Students learn the forms and their definitions (meanings).
	5 mins		Match	- Individual - Students will match all the target words with their meanings.	- Students learn the forms (oral & written) and write their definitions (meanings).
	10 mins		Spell (in English) Audio: On & Slow	- Individual - Students will listen to the word in English provided with its Arabic translation and image. - They will be asked to write the English equivalent.	- Students learn about phonological awareness and spelling.
	10 mins	Study flashcards Automatic-play for (3 mins) Audio: On	Writing (writing the word forms in Arabic) Audio: On	- Individual - Teacher guides students to choose Writing mode on Quizlet and select Arabic. - When students write a word, they will automatically listen to the word in Arabic & English.	- Students learn the forms (oral & written) and write their definitions (meanings).
	5 mins		Writing (writing the word forms in English) Audio: On	- Individual - Students will match all the target words with their meanings.	- Students learn the forms (oral & written) and write their definitions (meanings).
	10 mins		Test	- Individual - Students will answer 5 written questions, 5 matching questions, 5 multiple choice questions & 5 True/False questions.	- Students revise words using different tasks.
Materials: PCs, headsets, overhead projector				**Avenue:** LAB 3	
Follow-up: Delayed post-test & Attitude Questionnaire one week later.				**Assessment:** Immediate post-test (30 mins)	