



# **The influence of implicit and explicit learning through reading in a school environment: EFL adolescents' L2 vocabulary acquisition**

Yuna Jeon

Department of Education  
University of Oxford  
St Hugh's College

MSc in Applied Linguistics and Second Language Acquisition.

14<sup>th</sup> August 2024

Note that some graphs/tables/images may be removed in order to comply with copyright restrictions.

# DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION



1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [[https://ora.ox.ac.uk/terms\\_of\\_use](https://ora.ox.ac.uk/terms_of_use)].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [[https://ora.ox.ac.uk/deposit\\_agreements](https://ora.ox.ac.uk/deposit_agreements)] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Yuna Jeon
Date:	14 <sup>th</sup> August 2024

## TABLE OF CONTENTS

<b>Acknowledgements</b> .....	<b>4</b>
<b>Abstract</b> .....	<b>5</b>
<b>List of Tables</b> .....	<b>6</b>
<b>List of Figures</b> .....	<b>7</b>
<b>1. Introduction</b> .....	<b>8</b>
<b>2. Literature Review</b> .....	<b>11</b>
2.1. Implicit and Explicit Learning.....	11
2.2. L2 vocabulary learning through reading.....	16
2.3. Using glossaries for vocabulary learning.....	17
2.4. Theoretical background.....	18
2.5. The role of proficiency.....	20
2.6. Research Gap and Research Questions.....	20
<b>3. Methodology</b> .....	<b>22</b>
3.1. Research design.....	22
3.2. Participants.....	22
3.3. Instruments.....	23
3.3.1 Pre-test.....	23
3.3.2. LexTALE test for competent youths.....	23
3.3.3. The reading text and target words.....	26
3.4. Procedures.....	28
<b>4. Results</b> .....	<b>30</b>
4.1. Research Questions 1-2.....	30
4.2. RQ3. Is proficiency related to the student's performance?.....	35
4.3. Individual Item Level Analysis.....	36
<b>5. Discussion and Limitations</b> .....	<b>39</b>
5.1. Discussion.....	39
5.2. Limitations.....	41

5.2.1. The validity of the experiment environment.....	41
5.2.2. Unnatural reading environments.....	42
5.2.3. The lack of repetition and insufficient exposure.....	43
5.2.4. The adequacy of the explicit stimuli.....	44
5.2.5 Receptive and Productive Knowledge.....	45
5.2.6 Target words.....	45
<b>6. Conclusion.....</b>	<b>48</b>
<b>7. References.....</b>	<b>50</b>
Appendix 1: CUREC Approval.....	51
Appendix 2: Opt-out form.....	52
Appendix 3: Explicit and Implicit Texts.....	53
Appendix 4: Full reading text.....	54
Appendix 5. Target words glossary (A: adjective, V: verb, N: noun).....	59
Appendix 6. Statistics tables and charts.....	60
6.1. Histograms: test score distribution.....	60
6.2. Shapiro-Wilk Test: Q-Q Plots.....	61
6.3. Independent samples t-test results.....	64
6.4. A 2X2 mixed ANOVA.....	67
6.5. ANCOVA.....	71
6.6 Individual item level analysis.....	73

## Acknowledgements

Pursuing a master's degree was a long-held dream of mine, reignited by my students, Jaewoong Kim and Sihoon Kim. Returning to academia a decade after my undergraduate studies was challenging, and completing my dissertation would have been impossible without the support of many.

First and foremost, I would like to thank my supervisor, Dr. Faidra Faitaki, a professional and respected researcher and lecturer in the Department of Education at the University of Oxford. She has consistently guided me by providing detailed, critical feedback and warm support. Her research, intelligence, and kindness have greatly motivated me throughout my academic journey.

I am also grateful to my undergraduate professors, Prof. Jaeho Hwang and Prof. Bongchul Kim, who supported my efforts and served as referees for my master's application. Additionally, I would like to express my gratitude to Victor Safar, who helped me start my career both in the company and in graduate school.

It would be remiss not to mention my coursemates and friends in Oxford, who supported me throughout the year. Thanks to these brilliant and incredibly clever individuals, my knowledge has expanded far beyond what I had known for many years. Listening to topics like time travel paradox and philosophical issues unrelated to my dissertation over beers in pubs was undoubtedly one of the great pleasures of being in Oxford. Thanks also to friends from my undergraduate years, like So-Young Choi, for their enormous support and encouragement.

Above all, this academic journey would have been impossible without my beloved family. I cannot thank you enough for your trust and limitless support throughout my life. My mother, Doh-Im Hong, never lost faith in me, even when I doubted myself. My father, Yong-Sung Jeon, completed his PhD in his 40s while working and supporting a family with two young children. I respect both of you for your dedication to our entire family. My sister, Mina Jeon, my best friend, nemesis, ally, and advisor, knows me better than anyone. We have come so far and will only continue to thrive! I will always be there for you.

Finally, I dedicate this dissertation to my late grandparents, Insun Hwang and Jeongwon Hong, who always lavished unconditional, endless love on me. I love you, always.

## Abstract

This quantitative experimental study aimed to compare the relative effectiveness of implicit and explicit learning via reading on vocabulary acquisition and retention among English as a Foreign Language (EFL) secondary and high school Korean students. The study is based on Laufer & Hulstijn's (2001) Involvement Load Theory (henceforth, ILH), which was built to serve the purpose of encouraging theoretical and empirical research in second language (L2) vocabulary learning by proposing a construct of involvement, encompassing motivational and cognitive dimensions such as need, search, and evaluation. One hundred secondary and high school students participated in a 50-minute instructional treatment. Children in the explicit group read a text with a glossary, while children in the implicit group were only provided with the same text but without any glossary.

According to Nation (2006), learners must know the most common 8,000-9,000 word families to independently understand 98% of the content in various authentic novels or newspapers. Thirty-one target words were selected based on two criteria: frequency and familiarity. Two renowned word lists, the General Service List (GSL), 2,000 frequent words by West (1953) and the 8,000 British National Corpus (BNC), were used to validate the target word lists. Target words were not included in English's most 2,000 frequent words. Five out of thirty-one target words overlapped with the BNC and ranged between the most frequent 2,000 and 8,000 words. The remaining 26 target words were low-frequency words, beyond the first 8,000 families (Nation, 2006; see also: Schmitt & Schmitt, 2014).

These target words were assessed in both the pre-test and the post-test. The pre-test, which tested participants' prior knowledge of the target words, was a multiple-choice task that required learners to choose the correct definition of the 31 words. The post-test measured the retention of the target words after reading, using a cloze-test that required learners to fill in the blanks. Along with the two tests, LexTALE was conducted between the pre-test and the reading task to measure the students' English proficiency. After completing the reading task, participants were asked to finish the comprehension test immediately to check their attention towards the reading task.

The results showed a non-significant difference between the scores of the explicit and the implicit groups in the post-test. The relationship between language proficiency test scores, the pre-test scores, and the post-test results showed significant differences, meaning that prior knowledge and original proficiency are efficient predictors of vocabulary learning.

Finally, the research concludes by addressing discussions, limitations, future directions and implications of the study.

**List of Tables****Chapter 3. Methodology**

Table 1. Description of the Research Participants

**Chapter 4. Results**

Table 2. Test of normality: Shapiro-Wilk Test for each group

Table 3. Differences Between Explicit Group vs. Implicit Group on Three Vocabulary Tests

Table 4. Two-way mixed ANOVA: Within-subjects, Between-subjects

Table 5. Analysis of covariance summary table for post-test results with controlled variables

Table 6. Summary table of vocabulary learning outcomes per word (%)

Table 7. Summary table of vocabulary learning outcomes per student (word counts)

Table 8. Eight contextually-adapted words (non-literal meanings, word counts)

**List of Figures****Chapter 3. Methodology**

Figure1. The Sequence of the Experiment

## 1. Introduction

Vocabulary knowledge is one of the key predictors of learners' reading comprehension abilities, making it a crucial component of mastering L2 (Schmitt, 2008). Reflecting this importance, many researchers in second language acquisition have explored various facets of vocabulary knowledge through reading for L2 vocabulary learners (Ellis, 2015; Laufer & Hulstijn, 2001; Murphy, 2014; Nation, 2022; Schmitt & Schmitt, 2020).

Extensive reading is undoubtedly the most widely used and efficient tool in L2 vocabulary learning. At the same time, extensive reading introduces learners to large amounts of linguistic assets via massive quantities. It is pedagogically efficient as two activities, vocabulary learning and reading, happen simultaneously. Also, it opts for EFL learners, the only practical source to explore outside the classroom context (Grabe & Stoller, 2019; Huckin & Coady, 1999; Nation, 2022; Pigada & Schmitt, 2006). Krashen and his colleagues have strongly advocated for incidental vocabulary acquisition through reading, based on the theoretical hypothesis that vocabulary is learned in the same fundamental way as other aspects of language (Cho & Krashen, 1994; S. Krashen, 1989, 2013; Lao & Krashen, 2000; Mason & Krashen, 1997). The advantages of implicit learning not only lead learners to learn vocabulary but also aid them in acquiring additional aspects of vocabulary knowledge, such as grammar, multi-meaning words (polysemous words), and collocation via context. (Webb, 2007)

Of course, implicit learning is not a complete, holistic approach to vocabulary learning. Laufer (2003) claimed that explicit learning induces better learning because learners attend more intentionally to the words. Also, the author criticises implicit learning as too time-consuming and risky for learners. It might navigate learners through erroneous inducing procedures, which will result in the wrong vocabulary knowledge. Some researchers (Kweon & Kim, 2008; Min, 2008) addressed the low retention rate of implicit vocabulary learning is attributed to the failure of research methodology, mainly due to the amount of reading text, the number of target words or the kinds of text used.

Advocates of explicit learning argue that vocabulary retention and learning are more effective and faster when vocabulary learning is intentional (Schmitt, 2008). Additionally, explicit learning studies are often conducted alongside implicit learning, using the implicit learning condition as a control or comparison group. For example, Paribakht and Wesche (1996) conducted a classroom experiment comparing incidental learning with explicit instruction through reading among ESL university students. In their study, the enhanced group received reading text and vocabulary instruction, while the unenhanced group was provided only with the reading texts. The study's findings demonstrated that both conditions

cultivated significant vocabulary gains, with the reading-plus instruction group outperforming the reading-only group. However, a potential drawback of explicit learning could be that it may cause learners to become overly focused on memorising target words, leading to rote memorisation without engaging in active cognitive engagement that is necessary for 'learning' (Ahmad, 2012).

The underlying theoretical frame of this study is the Involvement Load Hypothesis (ILH), suggested by Hulstijn and Laufer (2001) and empirically proven by fellow researchers multiple times (e.g., Keating, 2008; Sung, 2013). The theory suggests that unintentional activities that provoke three elements- *need, search and evaluation* - out of unfamiliar words will result in greater vocabulary acquisition than activities that do not prompt these processes. The 'involvement load' concept includes motivational aspects (*need*) and cognitive elements (*search and evaluation*). Simply put, the theory claims the heavier the task load is, the better retention of vocabulary learners would yield.

Therefore, the discussion on implicit and explicit learning through reading is framed within the theoretical assumption that extensive reading facilitates vocabulary acquisition and that the explicit learning condition will lead to better vocabulary outcomes due to its inherently higher involvement load than the implicit learning condition. In simple terms, the implicit condition involves extensive reading materials, while the explicit condition includes additional stimuli designed to draw learners' attention to the vocabulary. Consequently, the hypothesis for this study assumes that learners in the explicit condition will demonstrate superior vocabulary acquisition compared to those in the implicit condition.

The literature reviewed thus far has predominantly focused on adult and child learners, often neglecting teenage learners in a crucial transitional phase. This oversight is particularly evident in studies involving Korean teenagers, where the roles of implicit and explicit learning through written text have received relatively little attention. Among studies relevant to Korean EFL learners and lexical or vocabulary knowledge (Bae, 2012; Choi et al., 2014; Chon et al., 2012; Jeon, 2007; Lee & Choi, 2020; Park et al., 2018; Park, 2001; Sung, 2013), very few empirical studies involved teenage EFL Korean learners. Jeon (2007) conducted a survey with 450 Korean high school students. In Bae's doctoral thesis (2012), the author recruited 262 near-puberty and teenage Korean students aged between 11 and 14. However, Jeon (2007) focused on vocabulary learning strategies such as rote learning and the relationship between learner's proficiency and learning strategies rather than investigating vocabulary learning through implicit and explicit learning. Likewise, Bae's (2012) study investigated the impact of English literacy instruction on the general literacy ability of learners and intercultural sensitivity.

In Lee and Choi's recent study (2020), two researchers investigated the effects of implicit and explicit instruction on morphological awareness and knowledge in 170 Korean high school students. The groups were divided into three: implicit learning, explicit learning, and a control group. The students were treated for six sessions with 6 derivatives and 24 items as target words. Participants took two pretests in inference and translation forms, and two delayed post-tests were conducted (one-week and seven-week). Scores were analysed by 3X2 mixed ANOVA. The results demonstrated that explicit learning was superior to implicit learning in all post-tests. Although implicit learning did lead to statistically significant gains in morphological awareness and knowledge, the effect sizes indicated only small to medium impacts. Although closely mirroring the design of Lee and Choi (2020), this study broadens the focus to encompass general vocabulary learning. Additionally, Lee and Choi's study concentrated on specific word instruction rather than embedding vocabulary within authentic reading materials.

Hence, this study investigates the effects of implicit and explicit learning via authentic reading on Korean EFL learners by comparing two learning conditions across randomly assigned participant groups. The literature review section explores a range of studies on implicit and explicit learning, with a focus on L2 vocabulary acquisition through reading. Among the studies that employed extensive reading as an intervention, those that provided glossaries for target vocabulary will be discussed in the context of the Involvement Load Hypothesis, which serves as the theoretical foundation for this study. Additionally, the role of language proficiency will be examined, as this study also aims to understand how learners' proficiency levels influence vocabulary acquisition. The study's research questions are outlined after identifying the research gap. Section 3, Methodology, details the research design, participants, instruments used in the experiment, and procedures followed. Section 4 presents the results, addressing the research questions and hypothesis. Section 5 discusses the findings, including the study's limitations. Finally, the conclusion wraps up the paper by considering future research directions.

## 2. Literature Review

In constructing the research narrative of this study, section 2.1 explores two learning conditions—implicit and explicit learning—posited within empirical research, and it includes comparisons between them. Section 2.2 explores the effect of L2 vocabulary learning through reading, which has been traditionally and widely used for L2 language learners. Section 2.3 examines the impact of vocabulary learning using glossaries. Discussions throughout sections 2.1 to 2.3 will be situated in the broader theoretical framework, the ILH. After addressing the role of proficiency, the research gap and research questions will be introduced. The study assumes three hypotheses for each research question.

### 2.1. Implicit and Explicit Learning

One of the long-debated subjects in L2 vocabulary acquisition that is better for L2 language learners is implicit or explicit learning, which displays contradictory outcomes in multiple empirical research studies. Implicit learning occurs without learners involving their intentionality or awareness, and explicit learning involves a conscious and intentional process. (Ellis et al., 2009) This section will present and analyse a few selected studies under both learning conditions.

One point that needs to be addressed beforehand is the terms used interchangeably below: explicit/intentional and implicit/incidental learning. Some studies strictly distinguish the terms, as they sometimes serve different purposes. The distinction is clear in separating L1 and L2 acquisition. While L1 acquisition happens implicitly, the L2 acquisition happens incidentally. Intentional learning indicates that the learners are informed they will be tested on target words. In contrast, incidental learning indicates the opposite, meaning the learners are unaware of a post-test. However, as Hulstijn (2003) pointed out, the above terms are not particularly relevant to vocabulary learning studies. More important is the quality of the mental processing during learning. Also, the two pairs similarly indicate that some language aspects are absorbed without intending to learn (L1), take instructions, or consult dictionaries (L2) (Dechert & Raupach, 1989). In this paper, any learning not involving explicit instruction is termed implicit learning. Conversely, any learning that involves making participants cognitively aware of what they are learning is considered explicit learning. (Nation, 2022)

#### Implicit learning

Deeply rooted in the distinction between acquisition and learning (Krashen, 1981), implicit learning has been studied by separating explicit learning based on the notion that formal language instruction involving conscious ‘learning’ is distinct from informal language

'acquisition'. Since then, many researchers tried to provide empirical evidence to support implicit learning through reading. Extensive reading is particularly useful for vocabulary growth and is called incidental learning. (Joseph & Nation, 2018; Nagy et al., 1985, 1987; Nation, 2022) They claim that new words are best learned when presented within texts, requiring learners to infer their meanings from the context. This perspective is founded on two key assumptions. First, learners invest more mental effort when they must infer or deduce solutions to problems than when they are simply provided with solutions. Second, information acquired through greater mental effort is more easily retrieved and recalled later than information gained with less effort. This idea, known as the *mental effort hypothesis* (Hulstijn, 1992), predicts higher retention for word meanings inferred from context than those given directly.

In another study that examined the effect of implicit vocabulary learning, Hulstijn (1996) experimented with Dutch students learning French. Participants were asked to read a French short story under one of three conditions: Marginal Glosses (providing L1 translations of unknown words), Dictionary (allowing the use of a bilingual dictionary), or Control (no additional support). After reading, the students were tested on their recall of 16 words that appeared once or thrice in the text. The study found evidence supporting the hypothesis that frequency of occurrence enhances incidental vocabulary learning, particularly when advanced L2 learners are given access to the meanings of unknown words through marginal glosses or by consulting a dictionary. In such cases, the repeated appearance of a word strengthens the connection between its form and meaning in the reader's mental lexicon. Conversely, when no external support is provided, learners are more likely to ignore unfamiliar words or incorrectly guess their meanings, diminishing the impact of word frequency.

Cho and Krashen (1994) conducted a small-scale empirical study with four native Korean adults who had lived in the United States for five years. Participants were asked to engage in recreational reading during their free time without specific requirements regarding the amount or type of text to read. A unique aspect of this study was that researchers did not impose any restrictions on dictionary use; participants were neither required nor prohibited from using a dictionary. Ultimately, only two participants used a dictionary throughout the reading treatment. Unlike other studies conducted in classroom or laboratory settings, this study is one of the few that placed participants in a natural reading environment. The total number of words read varied among participants. Three participants were tested orally and asked to define unknown English words they encountered during reading—words they had been encouraged to underline. The fourth participant was also tested orally but did not mark

unknown words during reading. Credit was given only when the vocabulary definitions were fully accurate.

The results showed that the two participants who used a dictionary gained an average of 17-37 words per volume, while the other two who barely used the dictionary or did not use the dictionary at all gained only seven to eight words per volume. This study suggests that learning vocabulary through natural reading alone may be slower than a combined approach that includes additional vocabulary learning strategies. It also indicates that integrating reading with supplementary learning activities can enhance the effectiveness of vocabulary acquisition.

The studies above indicate that implicit learning is effective only when supported by external factors such as repetition and dictionary use. Furthermore, DeKeyser (2003) concluded that limited evidence supports implicit learning based on a review of relevant L2 studies. The systematic review highlighted several disadvantages of implicit learning, particularly the lack of contextual clues, which can lead learners to infer incorrect word meanings. Inference, while useful, carries inherent risks as learners may easily make incorrect guesses, leading to the acquisition of inaccurate vocabulary. Moreover, successful inference often requires strong problem-solving skills, which not all learners possess. Brownfoot (2020) also raised concerns that young L2 learners might mistakenly believe they 'know' a word, only to realise they do not when attempting to recall its definition. Consequently, the effectiveness of explicit learning conditions has become more prominent for L2 learners. Given these limitations of implicit learning, it is essential to explore how explicit learning strategies can address these challenges and enhance vocabulary acquisition in L2 learners.

### Explicit learning

Explicit learning provides learners with direct cues or signals about the learning objectives. Scholars such as Beck & McKeown (2007) and Laufer (2003) have demonstrated the superiority of explicit learning conditions over implicit ones.

Beck and McKeown (2007) found meaningful outcomes in that the direct-instructed group received better learning results. The study taught advanced vocabulary to young children during their regular classroom hours, followed by a picture-based vocabulary test and verbal tests to assess learning. Considering the young learners, instruction was conducted in speech using the new research project, Text Talk, which provides rich instructions with target items in speech. Only the experiment group received the special treatment, while the control group only received the regular school read-out curriculums. Two

studies were carried out: Study 1 compared a directly taught group to a no-instruction group, while Study 2 compared the outcomes of a 6-day instruction group versus a 3-day instruction group. Both studies revealed that vocabulary gains were significantly higher in groups that received instruction, with the more extensively instructed group showing the greatest improvement. The implication is that the study gained empirical evidence that direct instruction also works in speaking conditions, and the rich and focused instruction facilitates the learners' huge improvement in learning advanced vocabulary, even for young learners. The age effect might differ, but the findings can also apply to teenage learners because older children or adults are more experienced and cognitively experienced.

Laufer (2003) critiqued the theoretical foundations of incidental learning and presented three experiments comparing reading with word-focused tasks. In Experiment 1, the retention of ten low-frequency words was compared between an implicit reading group and a group engaged in isolated sentence-writing tasks. The study involved 60 university-level EFL participants divided into two classes. Each class received a different task: the implicit reading group read a glossed text, while the explicit group wrote sentences using the ten target words without access to the text. Neither group was informed that they were participating in a vocabulary experiment until the post-test. The results favoured the sentence-writing group.

In Experiment 2, reading was compared to a composition task. This experiment involved 82 Israeli university EFL students, again divided into two classes. The same target words from Experiment 1 were used. The composition group received explanations in English and translations in Hebrew for the target words, after which they were asked to create their sentences using the ten words. The composition group showed higher recall scores than the reading group. Finally, in Experiment 3, three conditions were integrated: reading and consulting a dictionary, writing sentences with the ten words, and completing sentences provided by the researcher with the target words. This experiment involved 90 high school EFL learners with Arabic as their L1. Given the lower proficiency of these participants, the ten target words were replaced with less complex ones. Participants either looked up words in dictionaries, received Arabic translations for sentence writing, or were asked to complete sentences by filling in the blanks. The results again favoured the two groups engaged in sentence writing and completion, while the implicit reading group performed worse, even when they consulted the dictionary.

Although Laufer's study is valuable in that it directly compared implicit and explicit learning by employing various word-focused tasks and two different age groups (adults and teenagers), it is debatable whether glossed text represents an implicit learning condition. As

many other studies on glossed texts have shown, marginal glosses can act as explicit stimuli for vocabulary learners. Studies involving glossed texts will be further discussed in section 2.3.

### Implicit and explicit learning

Most studies comparing implicit and explicit learning suggest that explicit learning yields better than implicit learning. Previous studies also revealed that explicit instruction can enhance vocabulary acquisition by highlighting the relationship between form and meaning, while incidental vocabulary learning offers unique advantages. Incidental vocabulary learning allows learners to do reading and word learning simultaneously and enables them to gain more sophisticated knowledge of complex structures, such as collocations that require context knowledge. (MIRZAI, 2012).

Damhuis, Segers, and Verhoeven (2014) gave an example of each type of instruction. The study involved storybook reading with target words embedded but not explained (implicit; unintentional) and clear explanations of target words during storybook reading (explicit; intentional). The study examined the long-term effects of explicit versus implicit instruction on the breadth and depth of children's vocabulary. Two experimental groups learned two sets of 17 target words in order of explicit and implicit instruction. Their learning gains were compared to those of a 21-child control group without instruction. Both types of instruction led to broader vocabulary knowledge, but explicit instruction also resulted in deeper vocabulary understanding. These effects were maintained over time. Additionally, individual differences revealed that children with low short-term memory benefited more from implicit instruction in the short term, gaining greater vocabulary breadth. Still, they tended to forget more over time. In contrast, explicit instruction showed more sustained benefits across the board.

Dekeyser (2003) suggested an intriguing summarisation of multiple studies in classroom and laboratory environments: in both laboratory and classroom studies comparing implicit and explicit learning conditions, they consistently yield the same results that the explicit learning conditions are found to be more effective. However, the comparisons between implicit and explicit learning largely focus on reading with the reading-plus condition, an additional word-focused activity supplementing the reading-only condition. Only a few studies, including Laufer's (2003), address direct comparisons between implicit and explicit learning, and this current study intends to add one up and bridge the gap in the literature.

## 2.2. L2 vocabulary learning through reading

Reading is the major source of vocabulary acquisition in L2. Researchers have long studied using various reading materials such as textbooks, novels, or simplified texts. Many researchers have advocated extensive reading (ER) for incidental learning (Kweon & Kim, 2008; Mason & Krashen, 1997; Pigada & Schmitt, 2006; Suk, 2017), while Intensive reading (IR) is widely used in L2 classrooms, especially in EFL contexts (Park et al., 2018).

Kweon and Kim (2008) examined the effect of extensive reading on incidental vocabulary acquisition involving 12 Korean university students. Participants were instructed to read three novels over five weeks of the English courses and were tested on their vocabulary knowledge. A pre-test was conducted on the first day of the class, and an immediate and delayed post-test four weeks after the first post-test was conducted. Test scores were quantified for statistical analysis, demonstrating significant participant retention differences between the pre-test and the post-test 1. On the other hand, no significant differences were found between the first post-test and the delayed post-test. Researchers concluded that learners acquire vocabulary through ER, and the retention is maintained throughout four weeks (Hulstijn, 1992).

Lao and Krashen's (2000) study focused on Hong Kong EFL university students and their use of popular literature. Throughout 14 weeks, 91 students were asked to read six books, including one chosen by the students. All of the books were well-known novels, such as Jurassic Park. In addition to reading, the experimental group watched films related to four books, participated in class discussions after reading, and wrote two short essays. Meanwhile, the 39 students in the control group took a traditional academic skills course that covered the four language modalities: speaking, writing, listening, and reading. Vocabulary gains were assessed using the Accuracy Level Test in pre-and post-tests, and reading speed was measured with the Rate Level Test. The results showed that the experimental group significantly improved vocabulary knowledge and reading rate. As the authors noted, the experimental group was more proficient in English than the control group, consisting mainly of English translation majors. In contrast, the control group was composed of social science majors. Additionally, the experimental group was nearly three times the size of the control group. Despite these limitations, the analysis of gain scores revealed that extensive reading using authentic texts significantly improved learners' vocabulary.

In a quasi-experiment training study with Korean EFL university students, Suk (2017) demonstrated the potential of extensive reading. (Grabe & Stoller, 2019) The study was conducted with 171 participants learning English as a foreign language, dividing them into control and experimental groups. The 15-week reading courses involved regular textbook

use for the control group, while the experimental group also engaged in extensive reading for a portion of the class time. Both groups had similar out-of-class homework assignments, with the experimental group incorporating extensive reading of graded readers, contributing to their course grades. Assessments were conducted to measure reading comprehension, reading rate, and vocabulary gains before and after the treatment. The study found that the experimental group significantly outperformed the control group across all measures. The vocabulary gains were particularly notable, reflecting the effectiveness of the extensive reading component and the sensitivity of the testing instrument. The study's findings demonstrated significant improvements in reading comprehension and vocabulary learning for students who participated in the extensive reading program.

Research studies discussed above involved both L2 and EFL learners. In second language (L2) acquisition, terms such as foreign language and second language (L2) are distinguished. A foreign language is learned in an environment where it is not the major language of communication and is typically taught within a classroom setting in the learner's home country. Foreign language learners generally have limited exposure to their target languages outside the classroom and few opportunities to practice the language in real-life situations. Their motivation to learn the language stems from academic, professional, or personal interests. On the contrary, an L2 is acquired in a community where it is spoken alongside the learner's native language inside and outside the classroom. Learners are frequently exposed to the language in daily life. The motivation to learn the language often arises from necessity, such as for integration, employment, education, and social interaction within the community. (Murphy, 2014)

### **2.3. Using glossaries for vocabulary learning**

Glossing is an instructional method that provides meanings for selected words within a text. This technique draws explicit attention to lexical items during reading, where only incidental learning might occur (Schmitt, 2008). According to Nation (2022), glossing offers three key benefits: first, it allows readers to engage with texts that would otherwise be too difficult; second, it prevents readers from making incorrect guesses about the meanings of unfamiliar words; and lastly, it increases the likelihood that the glossed words will be learned. Glossed texts, or marginal glosses, have been widely employed in L2 studies to highlight target words and enhance the chances of these words being noticed and retained (Hulstijn et al., 1996; Ko, 1995; Lee et al., 2016; Lee & Lee, 2015).

The noticing hypothesis (Schmidt, 2001) contradicts Krashen's (1981) input theory that language is learned through subconsciously acquired comprehensible input. The noticing hypothesis states that learners must consciously observe and recognise the second

language's linguistic elements for acquisition to occur. In simpler terms, learners must focus on linguistic forms, even at a basic level. This idea lays the foundation for glossed text, highlighting the significance of deliberate learning in language acquisition.

Ko's (1995) thesis study investigated the effect of three different gloss conditions (no gloss, L1 gloss, and L2 gloss) along with the effect of incidental and intentional learning conditions over two months. Glosses were provided at the bottom of each page for 15 target words. The incidental group was informed that they would take a reading comprehension test, while the intentional group was aware of an upcoming vocabulary test. The study involved 189 university students divided into six groups and given three different reading materials. Following the reading, participants took a multiple-choice reading comprehension test and a vocabulary test, with a delayed post-test administered a week later without prior notice.

The results revealed no significant difference between the two learning conditions regarding vocabulary acquisition or reading comprehension. However, a significant difference was observed between the different gloss conditions on the post-test, with the L1 gloss condition showing the highest significance, followed by the L2 gloss and no gloss conditions. This finding that glossed text conditions outperform no-gloss conditions is also supported by other studies (Hulstijn et al., 1996; Watanabe, 1997).

#### **2.4. Theoretical background**

The Involvement Load Theory (ILT), articulated by Hulstijn and Laufer (2001), posits that the higher the involvement load in a learning task, the better the vocabulary retention. The level of engagement, known as "involvement load," suggests that when learners engage in activities that stimulate the need to understand, search for meaning, and evaluate unfamiliar words, they are more likely to expand their vocabulary than activities that don't require these processes. This concept is structured around a motivational-cognitive construct of involvement: motivational (*need*) and cognitive (*search and evaluation*) aspects. The concept of need has two levels: moderate and strong. *Need* is considered moderate when imposed by an external force, such as when a teacher assigns a task to students. On the other hand, strong *Need* refers to instances when students are internally driven, such as searching for a dictionary. *Research* and *evaluation* involve focusing on the relationship between form and meaning. *Research* is the process of trying to uncover the meaning of an unfamiliar L2 word or attempting to find the L2 word form that expresses a particular concept by referring to a dictionary or another authoritative source, such as a teacher. In the *evaluation* phase, learners compare the given word with others to determine whether it is

suitable for its context. *Evaluation* also entails making decisions about the meaning of a word by comparing all the meanings provided in a specific context and selecting the one that is the most appropriate. Hence, recognising the variation in several word meanings in a given context is described as moderate, while deciding how additional words will combine with the new word in an original sentence or text is described as strong evaluation. The combination of three factors, whether they are present or absent during the processing of a word in a natural or artificially designed task, constitutes the load of involvement. The level of prominence is represented by a factor marked as 0, moderate presence as 1, and strong presence as 2. The theory suggests that the more challenging the task, the better learners will remember new vocabulary.

To empirically validate their theory, Hulstijn and Laufer conducted two experiments with young adult EFL learners from Israel (n=128) and the Netherlands (n=97). Before the experiments, the participants took a pre-test to assess their familiarity with vocabulary. The researchers used three tasks with varying levels of involvement. Task 1 involved reading comprehension with L1 marginal glosses and a comprehension test. Task 2 involved reading comprehension with a "fill-in" task, a comprehension test, and a fill-in-the-blank test. Task 3 required the participants to write a letter incorporating the target words. The involvement index for each task was 1, 2, and 3, respectively. After completing the tasks, the participants underwent immediate and delayed (one-week) retention tests to measure their receptive knowledge of ten low-frequency target words. The analysis of the test scores revealed that tasks requiring a higher level of involvement, such as composition tasks, resulted in better retention compared to less demanding tasks like reading comprehension. Specifically, retention was highest in composition tasks, lower in reading plus fill-in exercises, and lowest in simple reading tasks. These findings suggest that more cognitively challenging tasks, which necessitate deeper processing of new words, lead to better retention. Therefore, the ILH underlines the importance of task load involvement in vocabulary learning, showing that higher involvement leads to improved retention of new vocabulary. The theory holds practical value not only for providing theoretical structure to L2 vocabulary researchers but also for pedagogical implications, such that teachers can consult this hypothesis when adjusting the level of involvement in line with their teaching goals.

Another study by Hulstijn (1996) found that words looked up in a dictionary during a reading task were better retained than those that were not or were only glossed in the text margin. Although marginal glosses did contribute to vocabulary retention, the strongest retention occurred in productive tasks, especially those that required using the words in an original context. This finding is supported by other studies showing that words practised in

productive tasks are remembered more effectively than those rehearsed in less active, unproductive tasks.(Cho & Krashen, 1994; Hulstijn, 1992; Keating, 2008; Sung, 2013; Watanabe, 1997)

## **2.5. The role of proficiency**

The relationship between vocabulary knowledge and language proficiency is reciprocal. Lexical knowledge not only serves as a powerful predictor of learners' language proficiency and academic achievement, but it also improves as learners' proficiency develops (Qian & Lin, 2019). This dynamic interplay is further supported by Xiao's (2015) study, which examined the impact of proficiency on the pragmatic competence of adult L2 learners. Xiao's findings demonstrated proficiency is a reliable indicator of overall L2 language competence. Although the study focused on general language ability, its conclusions provide a foundational framework for understanding how proficiency can be a legitimate predictor of L2 vocabulary acquisition. This reciprocal relationship has also been validated by other studies. For instance, Joseph and Nation (2018) found through their empirical research that children with stronger comprehension skills are more likely to become competent learners, emphasising the significance of proficiency in vocabulary learning. Additionally, Bensoussan and Laufer (1984) investigated the guessing ability of EFL students during reading tasks and discovered that a student's level of proficiency significantly influences their ability to infer word meanings. In other words, students with higher English proficiency can more accurately guess the meanings of unfamiliar words.

These studies collectively suggest that learners' original ability can be determined by proficiency in learning words before conducting the study. As learners' proficiency increases, so does their ability to navigate linguistic challenges, ultimately leading to more successful vocabulary acquisition.

## **2.6. Research Gap and Research Questions**

Although numerous studies on implicit and explicit vocabulary learning have been conducted with adult learners at the university or college level and young children, the importance of vocabulary learning among teenage learners has been largely neglected in the Korean context. It is almost surprising that studies comparing implicit and explicit learning for EFL learners in a classroom environment for Korean teenagers are virtually nonexistent.

Most research in these fields has focused on children or adults rather than teenagers. Similarly, within the context of Korean EFL learners, the majority of studies on vocabulary learning have been conducted with university students. This is likely due to the

fact that most researchers operate within higher education environments, where recruiting participants is less challenging.

This lack of focus on teenage learners is not unique to Korea. Researchers have also noted the scarcity of intervention studies in EFL contexts outside of Korea involving adolescent EFL learners. When they do exist, these studies often fail to have a meaningful pedagogical impact and contribute little to practical applications in education (Klingner & Vaughn, 2004; Townsend & Collins, 2009). Snow (2007) pointed out that most of the research in the volume centres on students from kindergarten through fifth grade, with a noticeable lack of attention to middle- and secondary-grade students. Consequently, critical issues that become increasingly important at these higher grade levels—such as vocabulary development, oral language proficiency, comprehension of complex texts, instruction on academic text structures, and the dynamics of student interactions—have not received adequate research attention.

Therefore, as a review of previous research has revealed scant empirical evidence of vocabulary learning studies involving Korean EFL teenage learners, this study aims to address this research gap by investigating whether implicit or explicit instruction affects vocabulary retention and comparing the effects of these two learning conditions. This research seeks to test the hypothesis posed by three research questions empirically.

- 1. Do implicit and explicit learning conditions facilitate Korean EFL adolescent learners' vocabulary learning while reading?**
- 2. If so, which condition (implicit vs explicit) is more effective?**
- 3. Is proficiency related to the student's performance/retention?**

For research questions 1 and 2, we assume hypothesis 1, that students are expected to learn the target words through reading. Hypothesis 2 predicts that the post-test scores will be higher in the explicit learning condition than in the implicit group based on the ILH. For research question 3, hypothesis 3 assumes that higher proficiency learners would yield better retention scores, as multiple studies have empirically supported it.

### 3. Methodology

#### 3.1. Research design

This research is a quantitative method study; an online experiment was determined to be the most effective approach for allowing South Korean EFL learners to participate and to yield quantitative data remotely.

#### 3.2. Participants

The project was approved by CUREC under the reference number **EDUC\_C1A\_24\_133** (see Appendix 1). 100 participants were recruited from four Korean schools: one middle school and three high schools. Other than the C high school participants, who are 18 years old and some of whom are preparing for the national college entrance exam, the rest are 17 years old.

Initially, the study aimed to recruit 16- or 17-year-old students (high school grades 1-2) to minimise age effects and maintain homogeneity within the participant group. However, it became unavoidable due to the limited number of schools willing to participate in the experiment, including a middle school. Moreover, this middle school is located in a neighbourhood known for its higher socioeconomic status and strong academic enthusiasm, distinguishing it from the other three high schools. Consequently, it was decided to include this school in the experiment and to treat all participants as homogeneous. Nonetheless, age could be a critical factor affecting the results; therefore, this oversight will be discussed further in the discussions section of the study.

Except for school A, where two classes participated, all other schools contributed one class each. The two classes of School A attended the online experiment sequentially. Further details are provided in Table 1 below.

**Table 1***Description of the Research Participants*

<b>School</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>Total</b>
Grade	2nd	2nd	3rd	3rd *	
Age	17	17	18	15	
Participants	35	21	17	27	100

\* D is a middle school, and A-C is a high school.

The recruitment procedure started by contacting school teachers interested in participating in the experiment. The teachers agreed to participate only if students and their parents agreed. Given the participants' age as competent youths and the nature of the experiment, an opt-out protocol was adopted. All students were informed to return the opt-out form if they did not want to participate in the study (see Appendix 2). No opt-out form was handed in.

### **3.3. Instruments**

This section sequentially presents the instruments used at each stage of the experiment. The sequence chart is available in Section 4.4. Procedure. (See Figure 1)

#### **3.3.1 Pre-test**

The first task is the pre-test, asking for definitions of 31 target words. It was presented in a multiple-choice form, giving three alternatives for the test-takers, excluding one correct option. They were given in the margin, with instructions on choosing the correct definition. The test measured how many target words the subjects had known before the reading intervention. One distractor was selected to be similar to the correct answer to avoid making the 31 items too easy or difficult. In contrast, the other was chosen to be irrelevant or different. This was necessary because if the alternatives were too plausible, there would be a risk that participants may infer the wrong meaning.

#### **3.3.2. LexTALE test for competent youths**

In psycholinguistics experiments, proficiency measurements serve two primary purposes: 1) to categorise participants into groups based on smaller or larger vocabulary sizes and 2) to exclude participants who fall below a specified proficiency threshold. (Lemhöfer & Broersma, 2012). This study employed the LexTALE test to measure participants' proficiency levels,

setting a threshold and using the scores to investigate the correlation between language proficiency and vocabulary acquisition.

LexTALE is a validated lexical test for intermediate to advanced adult English learners designed by Lemhöfer and Broersma in 2012. The primary purpose was to offer researchers a practical and objective measure of English proficiency instead of relying on self-ratings and language background questionnaires. The test evaluates the test takers' receptive lexical knowledge using 40 English words and 20 pseudowords. Notably, 40 English words are considered 'subjective nonwords', meaning the frequency of these 40 words is too low that participants are unlikely to recognise them easily. Therefore, to balance the proportion of the two categories, the valid English words are double the number of pseudowords. Test takers are directed to respond with either yes or no to determine if each item is a valid English word. The validity of the test was confirmed through external criteria such as translation tasks, proficiency tests such as TOEIC, and experimental data in lexical decision tasks. The results showed strong correlations between LexTALE scores and participant performance in these experimental contexts (Lemhöfer & Broersma, 2012).

However, the test is often criticised for its limitations. One major critique is that, like other tests, such as the Vocabulary Size Test by Nation and Beglar (2007), the target words from frequency bands are small, random samples. These tests select items from frequency bands under the assumption that items within the same band will exhibit similar difficulty levels (Beglar & Nation, 2007; Gyllstad et al., 2019; Stoeckel et al., 2021). However, empirical research has shown that frequency alone does not adequately predict word difficulty, accounting for only about 25% of the variation in word recognition difficulty (Hashimoto & Egbert, 2019; Qi et al., 2024). Consequently, items selected from the sample pool may not accurately represent their frequency band or reflect test takers' proficiency level. Moreover, a recent study has demonstrated that LexTALE has a relatively low correspondence to the L1 and proficiency level of participants, leading to recommendations for caution when using LexTALE in future studies. (Puig-Mayenco et al., 2023). Despite these concerns, many researchers agree that LexTALE is a quick and easy implementation, which makes it a widely used test for measuring receptive vocabulary in research settings. (Izura et al., 2014)

LexTALE has been rarely used to measure teenage learners' vocabulary proficiency. That is primarily because the test was aimed to measure intermediate or high-proficiency adult English learners. A recent study used LexTALE to test adolescents' vocabulary proficiency. Qi and colleagues (2024) developed a bespoke LexTALE to measure Chinese vocabulary proficiency. The LexCH tested 480 students from a junior high school and a high

school and found out this test was valid for L1 Chinese speakers. Since it has not yet been empirically validated for Chinese L2 learners, testing teenagers with LexTALE lacks empirical support. To sum up, this means testing teenagers using LexTALE has yet to be done.

However, this test was chosen in the present study as the baseline of our participants' English ability. This choice was motivated by the fact that the participants were older teenagers (they would become adults age within 1-3 years) and had studied English since Grade 3 (age 9-10) for a minimum of 7 years. According to the recent study by Choi (2023), which compared Korea's achievement levels in the national English curriculum and high school textbooks to Common European Framework of Reference for Languages (CEFR) Levels, Korean high school students in a formal academic environment are expected to reach English proficiency levels ranging from B1 to B2. These levels correspond to intermediate proficiency on the CEFR, a globally recognised scale for measuring English proficiency. Therefore, it is reasonable to expect that participants in this study would have at least an intermediate level of English proficiency, and the use of LexTALE on competent youths has been justified.

In conclusion, English LexTALE was conducted following the pre-test. The test contained sixty words, including 20 unfamiliar words and 40 pseudo words that did not overlap with 31 target words excerpted from the intervention reading text, *The Giver*. Before starting the test, participants were instructed both in L1 (Korean) and L2 (English) to mark "✓" when they knew or were convinced that they were English words. For words they doubt or think do not exist in English, they were also instructed to mark "x". A 10-minute time limit was implemented for this test to ensure that the experiment could be completed smoothly within the allotted class period, as formal class sessions in Korea last 50 minutes. This constraint was necessary to effectively manage the overall experimental duration of a maximum of 50 minutes.

A two-minute break was first taken after the LexTALE test to give participants intermittent short breaks. After the break, participants were led to the main intervention: reading the authentic text. Each group was also instructed in both Korean and English to read the text within 15 minutes. For the explicit group, 31 target words were marked in bold throughout the reading to indicate they were glossed on the right side margin of the text. The meaning of the target words in the side margin was also described in English (L2). On the other hand, the implicit group received the plain text without boldfaced target words or glossaries, which is a No-Clue condition. (See experiment screen in Appendix 3)

After the reading session, participants were immediately instructed to take the

reading comprehension test. This test acted as an attention checker. In other words, it assessed whether the students paid appropriate attention to read the whole text. This was done to screen out the participants who didn't pay enough attention to the experiment. Suppose the student's comprehension rate is zero or one, but their retention rate of target words is higher than others. This can be interpreted as the student who knew the target words before the experiment or had a lucky guess for the correct answer without paying attention to the text. This test consisted of three simple Multiple-choice questions to examine whether students understood the novel's narrative and contents.

After another intermittent two-minute break after the comprehension test, participants were instructed to continue to the post-test. The questions on the test focused on students' ability to recall the vocabulary they read during the reading task. The test was in cloze form, measuring participants' productive vocabulary knowledge by asking them to input the correct word in the blank of the sentence. Approximately 26 % of the target words switched the original meanings to contextually adapted meanings to test if any learning occurred during the 15-minute intervention, regardless of implicit or explicit conditions. The contextually adapted meanings of the target words will be further explored in the reading text paragraph below. The pre-test and the post-test were designed to be similar except for the test formats and the two factors: First, the pre-test measured participants' receptive knowledge and the post-test measured productive knowledge. Second, the post-test included contextually adapted meanings of 8 words, the most distinctive difference between the pre-test and the post-test. Both tests consisted of multiple-choice questions, with 31 questions, each weighted one, and thus the possible score range was between 0 and 31.

### 3.3.3. The reading text and target words

The reading text that the intervention was based on was chapter 11 of "The Giver" by Lois Lowry. (See Appendix 4 for full text) This science-fiction, young adult novel was first published in 1993 and earned the Newbery prize the following year. The book is also included in the Honor List by Nilsen & Donelson (2009), who stated that young adult novels are "stylistically less complicated, which is a significant factor in language learning" (Kweon & Kim, 2008, p 195). The story's background is a dystopian world where citizens have no freedom. Also, people are not allowed to have memories except for the person known as the Receiver of Memory. The main story is about a boy named Jonas, the next Receiver, who, upon receiving the memories of the past, has an awakening about the society that he lives in. Chapter 11 is about Jonas receiving pleasant and painful memories from the previous Receiver, who shows the boy what snow sledge and sun basking are. This novel was selected as it was deemed interesting enough to draw participants' attention and maintain

their focus while they went through the reading intervention. Chapter 11 is 8 pages long (pp.80-87) and 1,978 words long.

31 target words were selected from Chapter 11 (see Appendix 5) and converted into root forms. Nagy and colleagues (1987) estimated that children learn 4,000 to 6,000 words from context while reading. Assuming the median value of this category at 5,000 words, dividing it over nine months excluding vacation periods, and considering that the average monthly instructional time for regular English classes in Korean schools is 20 hours, Korean students are estimated to learn approximately 27-28 words per class. Furthermore, accounting for the possibility that students already know some of these words, a final list of 31 words was selected for the study. The target words consisted of eight nouns, seven adjectives and sixteen verbs, with verbs two times higher than nouns. Usually, nouns are better learned than verbs for children because they are conceptually more straightforward (Gentner & Beranek, 1982). However, a recent study suggests unparalleled results to this assumption, claiming that the approach should be comprehensive regarding word learning, not the dichotomisation of nouns and verbs. Hirsh-Pasek and colleagues (2006) found that children learn faster for more concrete and shape-based words at an early age, but not the notion that 'nouns are learned before verbs'. Therefore, based on the recent research study, this study ultimately concluded that using 31 target words, half of them being verbs, would not be problematic for teenage participants.

After selecting the target words, the glossary was verified by using the two word lists that are frequently used in vocabulary research: *the General Service List (GSL)*, which has the 2,000 most frequent headwords (West, 1953) downloaded from the *Lextutor* website and the British National Corpus (BNC) list-*Rank frequency list for the whole corpus (not lemmatised)*. The target words were not included in the GSL, meaning that, frequency-wise, the target words are beyond 2,000 words. Only five words (***firmly, descent, abruptly, treat, consciousness***) overlapped with the BNC list, with the frequency rate (Frequency per million words) from 11 to 40, which means that readers are likely to encounter these words from 11 to 40 times out of a million words. Conclusively, these five target words are between 2,000 and 8,000 frequent words, and the rest of the 26 target words can be assumed to be low-frequency words beyond the 8,000 most frequent families (Schmitt & Schmitt, 2014), which this study assumed the target words are unfamiliar or unknown to teenage EFL students. Additionally, to measure if participants would pick up contextually adapted meanings while reading, eight words (***pepper, pinprick, treat, tug, firmly, crease, grasp, and drain***) among the target words were tested differently in the post-test to the pre-test. These eight words take approximately 26% of the total target words list.

Lastly, to determine the reading time's duration, the intervention text's word count was divided by the average reading rate per minute. According to Brysbaert (2019), for English silent reading, L1 readers read an average of 238 words per minute for non-fiction and read slightly faster (260 wpm) for fiction. Old or L2 readers were observed to read slower than L1 readers by 10%, which makes 234 wpm (Dirix et al., 2019, as cited in Brysbaert, 2019). Considering the reading task of this study is a total of 1,978 words, the participants would be expected to read the text within approximately 8.5 minutes. ( $1978 / 234 = 8.45$ ) Therefore, the time given (15 minutes) to the participants for the reading task was sufficient.

### 3.4. Procedures

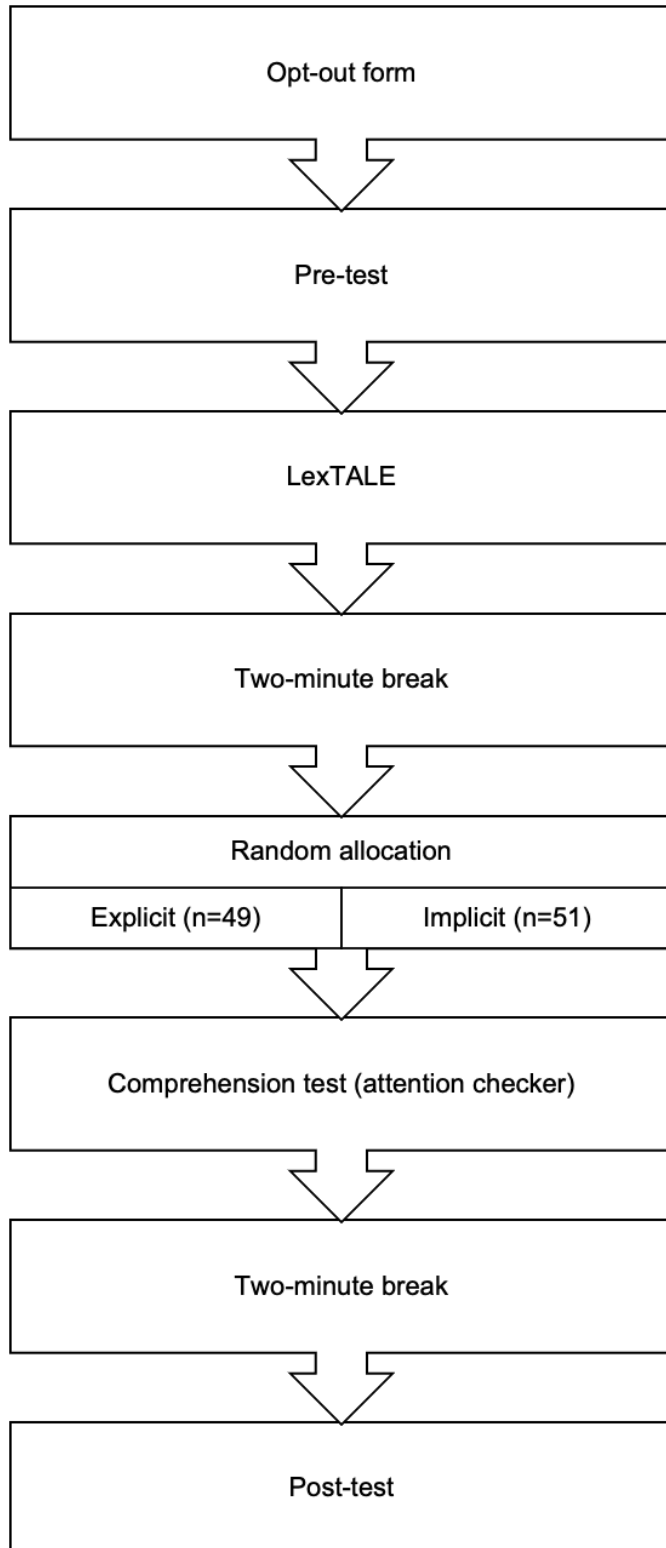
This study adopted a quantitative method design. The online experiment program (Gorilla) randomly allocated each subject that entered the experiment to the implicit or the explicit group. The procedure of the classroom experiment consisted of five steps: the *Pre-test (Multiple Choice)*, *English LexTALE*, *Intervention (Reading the novel excerpt)*, *Reading comprehension test (Multiple Choice, attention checker)*, and *Post-test (Cloze)*. Before the experiment, all participants were informed that there would be a pre-test and a post-test to create the intentional learning task environment, meaning if participants know that a test is coming up, learners are more likely to pay attention and focus on the task. (Hulstijn, 1992)

At the beginning of the experiment, the participants were asked to read a brief information sheet and check the consent form. They were then instructed to start the pre-test by clicking the 'next' button.

For scoring, LexTALE data was manually calculated using the best-score yielding method recommended by the inventors of the test: averaged % correct. It is calculated using the formula as follows:

$$((\text{number of words correct}/40*100) + (\text{number of nonwords correct}/20*100)) / 2$$

This ensures that a bias towards either yes or no responses is equally penalised, regardless of the differing numbers of words and nonwords. Before calculating the score, the first three items (*dogify*, *party*, *bargoin*) were excluded from the Excel data, as they are dummies. (classified as 'filler' words in the Excel data). These three dummies were embedded to inform test takers about how the questions would be presented and familiarise them with the test format. (Lemhöfer & Broersma, 2012).

**Figure 1***The Sequence of the Experiment*

## 4. Results

This section presents the results to address results questions 1-3. The data was downloaded in Excel from the online experiment platform. The files were processed to meet the proper format for SPSS.

### 4.1. Research Questions 1-2

The first research question was if two learning conditions facilitate Korean EFL adolescents' vocabulary learning while reading. For hypothesis 1, both groups were expected to learn words through reading. The second research question was to examine which condition (implicit vs explicit) was more effective in gaining words. For hypothesis 2, it was assumed that the explicit group would yield higher scores in the post-test for the target words they learned from reading than the implicit group.

The descriptive analysis was run to get an overview of the entire data. Table 1 provides insights into four test scores. The kurtosis and skewness showed a near-normal distribution, as both ranged from -1 to 1, except for the post-test kurtosis (-1.1). Participants knew more than half of the definitions ( $M=17.74$ ), while the mean score of the post-test scores was decreased by 3.04. LexTALE exhibited the highest variability in central tendency and dispersion, indicating participants' English proficiency (at least in vocabulary proficiency) is very diverse. For more statistical distributions, see Appendix 6.1 for histograms.

**Table 1**

*Descriptive Analysis of Experiment Tasks*

	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>SD</b>	<b>Skewness</b>	<b>Kurtosis</b>
Pre-test	7.00	29.00	17.74	5.27	<b>0.10</b>	<b>-0.67</b>
LexTale	21.00	86.25	53.79	11.76	<b>-0.27</b>	<b>0.53</b>
Comprehension	0.00	3.00	1.25	0.96	<b>0.36</b>	<b>-0.79</b>
Post-test	3.00	31.00	14.70	7.46	<b>0.43</b>	<b>-1.10</b>

*Note.* N = 100

**Table 2***Test of normality: Shapiro-Wilk Test for each group*

	Implicit		Explicit	
	Sig	Statistics	Sig	Statistics
Pre-test	0.695	0.984	0.338	0.974
LexTALE	0.148	0.966	0.604	0.981
Comprehension	<b>&lt;.001</b>	0.869	<b>&lt;.001</b>	0.866
Post-test	<b>0.007</b>	0.934	<b>&lt;.001</b>	0.902

The Shapiro-Wilk test was used to verify the normality of all measurements and decide whether to use a parametric or nonparametric test because the Shapiro-Wilk statistic (W) can handle sample sizes between 3 to 5,000 (Razali & Wah, 2011). The detailed W statistic results are in Table 2.

According to the test results, the significance value of the Shapiro-Wilk test is greater than 0.05 for the pre-test and LexTALE, meaning that the data is normally distributed for these two tests. However, the data is not normally distributed for the comprehension and post-test tests. ( $p < 0.05$ ). Hence, the comprehension test was eliminated from the analysis because not only did the data reject the null hypothesis that the sample comes from a normal distribution, but they were also used as an attention checker for the test participants. The not-normally distributed post-test scores were transformed using Log10 in SPSS, but again, they showed a *p-value* of less than 0.05. (For implicit group, 0.02, for explicit group, 0.03)

According to Kim and Park (2019), the Central Limit Theorem posits that “the distribution of sample mean values tends to follow the normal distribution regardless of the population distribution if the sample size is large enough” and that “...the central limit theorem guarantees the normal distribution of the sample mean values” (p. 333). The Shapiro-Wilk test is particularly effective for identifying non-normality in short-tailed or skewed distributions. The test is less effective with symmetric, moderately long-tailed distributions (Royston, 1992). Given that this study's data shows a slightly flattened distribution with a long tail on one side (the post-test kurtosis is -1.1), there is a possibility that the Shapiro-Wilk test may have overstated the deviation from normality. Consequently, it is reasonable to conclude that the data in this study are normally distributed. (Appendix 6.2)

**Table 3***Differences Between Explicit Group vs. Implicit Group on Three Vocabulary Tests*

	Implicit (n=51)		Explicit (n=49)		t (98)	p	Cohen's d
	M	SD	M	SD			
Pre-test	18.10	5.46	17.37	5.11	0.69	0.49	0.14
LexTALE	53.69	11.90	53.89	11.75	-0.08	0.94	-0.02
Post-test	15.73	7.77	13.63	7.06	1.41	0.16	0.28

*Note.* Levene's Test indicated that the homogeneity of variance assumption was not met for the variable. Hence, equal variances are assumed for all three tasks (Definition, LexTALE, and Cloze). For more statistics, see Appendix 6.3. Degrees of freedom are all equal (df=98).

Based on the assumption of data normality, independent t-tests were conducted on tasks that can assess whether there are significant differences in mean scores between the two groups, implicit and explicit learning. The null hypothesis is that the means of the two populations are equal. The independent variables are the implicit and explicit group, whereas the dependent variables are the pre-test, LexTALE, and post-test scores. The comprehension test, primarily serving as an attention checker, was excluded from this analysis due to its almost zero significance in the Shapiro-Wilks test (See Table 2).

The results indicated that the p-values for Levene's test for the three assessments exceeded the significance level  $\alpha = 0.05$ ; thus, the null hypothesis could not be rejected, and equal variances were assumed ( $p = 0.857$  for the pre-test,  $p = 0.796$  for LexTALE, and  $p = 0.316$  for the post-test). Table 3, which compares the mean scores of three different tasks across each learning condition group, shows no significant differences between the groups on the pre-test:  $t(98) = .69$ ,  $p = .49$ , with a small effect size ( $d = .14$ ). Similar results were observed for the post-test:  $t(98) = 1.41$ ,  $p = .16$ , suggesting no significant difference between the implicit and explicit groups, with a small effect size ( $d = .28$ ). The groups did not differ regarding their LexTALE scores either:  $t(98) = -.08$ ,  $p = .94$ ,  $d = -.02$ .

The pre-test scores indicated that the implicit group scored slightly higher ( $M = 18.10$ ,

SD = 5.46) than the explicit group (M = 17.31, SD = 5.11). For the LexTALE test, both groups demonstrated roughly equal means. The post-test scores, however, showed a more noticeable discrepancy between the means of the two groups, with the implicit group students averaging two more points in the post-test than the explicit group. However, the most notable thing is that the post-test scores for both groups decreased compared to the initial test. Statistically speaking, the figures mean that learning has decreased, contrary to the expectation of RQ1. However, we must be cautious when interpreting the results here because the two tests (pre-test and post-test) are different. They are not in the same format, and they test different vocabulary knowledge. (receptive vs productive) For more discussion, see Section 5.

A 2 X 2 mixed analysis of variances (ANOVA) was conducted to investigate whether gains in vocabulary knowledge over time differed by learning type. The between-subjects factor is the type of learning conditions (implicit or explicit), and time conditions (pre-test and post-test) are the within-subjects factor. The assumptions are met for ANOVA; the independent variables are categorical, and the dependent variables are continuous. This test was selected to estimate how the dependent variables (quantitative) change upon the independent variables.

**Table 4**

*Two-way mixed ANOVA: Within-subjects*

Source	Time	F (1,98)	$\eta^2$	Sig.
Time	post-test vs. pre-test	19.36	0.17	<0.001
Time * Groups	post-test vs. pre-test	0.96	0.01	0.33

*Two-way mixed ANOVA: Between-subjects*

Source	F (1,98)	$\eta^2$	Sig.
Groups	1.69	0.02	0.20

The F-value for the main effect of time indicates that the difference between the pre-test and the post-test is statistically significant [ $F(1,98) = 19.36, p = <.001, \eta^2=0.17$ ] The interaction between time and group ID is not significant, suggesting that the effect of time does not significantly differ across different groups. [ $F(1,98) = 0.96, p = 0.33, \eta^2 = 0.01$ ] For the tests of between-subjects effects, there was a non-significant effect of learning conditions on vocabulary knowledge [ $F(1,98) = 1.69, p = 0.2, \eta^2 = 0.02$ ]. The visual inspection of the profile plots also indicated that both groups started not significantly different from each other at the initial test (pre-test) but got significantly different at time 2 (post-test). (See Appendix 6.4)

In conclusion, the mean values indicate that learning decreased, unlike the expectation that reading would yield vocabulary learning for both groups. According to the ANOVA result, neither implicit nor explicit learning conditions impacted the students' vocabulary learning. There are no significant differences between the two groups, meaning that no particular learning condition had a meaningful impact on the three test scores.

#### 4.2. RQ3. Is proficiency related to the student's performance?

To address RQ3, a between-group analysis of covariance (ANCOVA) was conducted to examine if learners' proficiency levels affected vocabulary retention performance. Two covariates were the pre-test and LexTALE scores. The independent variables were the treatment conditions: explicit and implicit, labelled as 0 and 1, respectively (0 for the implicit group and 1 for the explicit group), whereas the dependent variable was the post-test score. ANCOVA provides a clearer picture of the actual impact of the independent variables on the dependent variable while controlling the effect of extraneous variables. By including the pre-test and LexTALE scores as covariates, ANCOVA adjusts the post-test scores to account for initial differences in vocabulary knowledge and baseline performance levels. This adjustment helps to isolate the effect of the learning method itself on vocabulary retention. The results obtained are in Table 5.

**Table 5**

*Analysis of covariance summary table for post-test results with controlled variables*

Source	SS	df	MS	F	p	$\eta^2$
Pre-test	600.84	1.00	600.84	14.49	<.001 *	0.13
LexTALE	351.45	1.00	351.45	8.47	0.004 *	0.08
Groups	77.14	1.00	77.14	1.86	0.18	0.02
Error	3981.70	96.00	41.48			
Total	27126.00	100.00				

*Note.* SS sum of squares, MS mean square, \* $p < .05$

According to Levene's test result, the null hypothesis of equal error variances is not rejected. [ $F(1,98) = 0.76, p = 0.38$ ] The data meets the homogeneity of variances assumption. The overall corrected model is highly significant ( $F = 12.339, p < 0.001$ ), meaning that the independent variables account for many differences in post-test scores among participants. In other words, these predictors are strongly related to how well participants performed on the post-test. (See Appendix 6.5) The model explains 27.8% of the variance in the post-test scores, as indicated by the R squared value (adjusted R squared = 25.6%). The first covariate, the pre-test scores, is a significant predictor ( $p < 0.001$ ) with a small effect size ( $\eta^2 = 0.131$ ). This explains the role of prior knowledge in

vocabulary learning. The other covariate, LexTALE scores, significantly predict the post-test scores ( $p = 0.004$ ), with a small effect size ( $\eta p^2 = 0.081$ ). This indicates that vocabulary knowledge, as measured by LexTALE, significantly influences post-test performance.

On the contrary, the learning conditions do not significantly predict the post-test scores when controlling for pre-test and LexTALE scores, with an F-value of 1.860 and a significance level of 0.176. The partial eta squared value of 0.019 suggests a very small effect on the post-test scores by the intervention groups (implicit vs. explicit).

In conclusion, the ANCOVA results demonstrate that while the overall model is significant, and covariates like LexTALE and pre-test scores significantly predict post-test performance, the learning conditions do not substantially impact vocabulary learning when these covariates are controlled for. The significant covariates suggest that vocabulary knowledge and prior performance are meaningful in learning vocabulary. However, the type of vocabulary learning conditions does not lead to significantly different test scores. The findings imply that vocabulary learning happens better when learners hold higher language proficiency regardless of the learning conditions. Hence, hypothesis 3 is supported by statistical data, indicating proficiency levels and pre-existing knowledge are predictors of the post-test (with small effect sizes) but not the learning conditions.

### **4.3. Individual Item Level Analysis**

An individual item-level analysis was manually conducted to determine how many words participants learned and what words they learned via the reading intervention. This analysis simply calculated the integrated data of the pre-test and the post-test scores. A table listing 31 target words was constructed in a single column, with each row representing a participant's ID and their corresponding pre-test and post-test results arranged horizontally. Then, the results were categorised as follows: when a participant answered incorrectly on the pre-test but correctly on the post-test, it marked "O", indicating learning; "=" indicated existing knowledge (prior knowledge), where participants responded correctly on both tests; and "X" indicated no learning, applied to both cases of Correct-Error and Error-Error responses.

**Table 6***Summary table of vocabulary learning outcomes per word (%)*

Per word	Average %	Maximum %	Minimum %
Learned	17.20%	51.6%	0.0%
Not-learned	52.60%	90.3%	0.0%
Prior knowledge	30.20%	93.5%	3.2%

**Table 7***Summary table of vocabulary learning outcomes per student (word counts)*

Word	Average	Maximum	Minimum
Learned	17.23	32	7
Not-learned	52.58	76	40
Prior knowledge	30.19	49	12

\*For more statistical data, see Appendix 6.6.

The results are detailed in Tables 6 and 7. Table 6 summarises the percentage of vocabulary learned, not learned, and known before the intervention, evaluated per word. On average, students learned 17.2% of the target vocabulary, equating to five or six words. Notably, the participants did not acquire 52.6% of the target words—approximately 16 words. On average, students had prior knowledge of 30.2% of the vocabulary or about nine words. The highest acquisition percentage recorded was 51.6%, indicating that the most successful student learned over 15 target words through reading. In contrast, the least successful student did not acquire 90.3% of the target words. Remarkably, one student already knew more than 90% of the target words before the intervention.

Table 7 presents the per-word acquisition counts. The word 'bask' was learned by the most students (32), whereas 'pepper' was the most challenging word, with 76 students failing to acquire it. The word 'comprehend' was the most widely known before the experiment, with nearly half of the participants already familiar with it. Both Tables 6 and 7 highlight the challenges of vocabulary acquisition in this study, as evidenced by the low overall rates of vocabulary growth.

**Table 8***Eight contextually-adapted words (non-literal meanings, word counts)*

words	O	X	=	Literal	Non-literal
<b>Crease</b>	29	59	12	A line or fold in fabric or paper	Deep wrinkles on the face
<b>Drain</b>	15	51	34	To remove or draw off liquid	Consume all the energy, exhausted
<b>Firmly</b>	27	44	29	Tightly	With determination and strength of character.
<b>Grasp</b>	16	45	39	To hold something	Pick up the new concept
<b>Pepper</b>	7	76	17	A spice derived from ground peppercorns	Scattered
<b>Pinprick</b>	11	49	40	A very small puncture or prick	Stinging pain
<b>Treat</b>	10	63	27	To provide or offer something enjoyable	Reward, Gift
<b>Tug</b>	12	58	30	To pull with force	Recall something from the past

Learned: O / Not-learned: X / Prior Knowledge: =

An analysis of the eight words with non-literal meanings was conducted to determine whether they were particularly challenging to learn, given that they share the same context but have different meanings in the pre-test and post-test. The results are presented in Table 8. The data suggest that non-literal meanings were generally more difficult to learn, with the exceptions of "Crease" and "Firmly" when compared to the overall averages observed in Table 7 ( $M = 17.23$ ). The two words with the highest number of participants failing to learn them were "Pepper" and "Treat." This finding is particularly interesting because the word class of these two items changed when used in different contexts. Specifically, the literal meaning of "Pepper" was a noun in the pre-test, but in the post-test, the non-literal meaning required participants to understand it as a verb. Similarly, "Treat" shifted from its literal meaning as a verb (to provide or offer something enjoyable) to a noun in its non-literal sense (meaning reward or gift). This analysis suggests that even within non-literal meaning words, a change in the part of speech may impact vocabulary acquisition more. In other words, learning multiple meanings of a word is challenging and may require more cognitive engagement than learning the literal meanings alone. This will be further elaborated in the discussion section.

## 5. Discussion and Limitations

### 5.1. Discussion

This research investigated vocabulary learning facilitated under two different learning conditions in Korean EFL teenage learners. The age of learners should be addressed first. The inclusion of middle school students was not validated. The age factor could have been treated by implementing a pilot study, but we could not launch one due to the difficulty of recruiting more participants. Also, participants' school information was all anonymised in the data and became untraceable. Therefore it is impossible to separate the middle school students' data alone, resulting in the study not knowing how this inclusion of young learners affected the results. Even though the socioeconomic background is high, there is no guarantee that all groups share similar proficiency. LexTALE results are similar across the groups, and there is no visible outlier in LexTALE Q-Q plots (See Appendix 6). Therefore, again, we assume these young learners' English ability is similar to the rest's.

Research questions #1 and #2 explored whether implicit or explicit learning leads to vocabulary learning and, if so, which condition is more effective. Based on the ILH, the predicted outcome was that the retention rates would be higher in the explicit learning condition, as it includes one additional load factor- the glossed text- than in the implicit condition. The results revealed unexpected findings, but they were interesting enough to discuss. The statistical analysis indicated that vocabulary knowledge decreased over time.

This decrease may be attributed to the possibility that the provision of glosses might made the task visually and cognitively more challenging for the learners (see Appendix 3). The definitions' density in the marginal glossaries may have been overwhelming. Additionally, while learning through reading in one's L1 is widely supported by researchers (Nagy et al., 1987) and is generally endorsed in extensive reading studies of L2 acquisition (Hulstijn et al., 1996), it is important to note that the process is more demanding in L2. L2 learners often lack the necessary vocabulary to find contextual clues facilitating word learning. In EFL settings, if reading is indeed the primary method of instructed vocabulary acquisition, then rather than focusing on isolated word practice, learners should be encouraged to read as extensively as possible, both inside and outside the classroom, similar to the approach taken with L1 learners (Laufer, 2003). The limitations discussed in the following section delve into these findings, which contradict the original research hypothesis.

Research question #3 examined whether learners' English proficiency levels and pre-existing knowledge of target words impacted post-test scores. The hypothesis was that

higher-proficiency learners would achieve better results than lower-proficiency learners. The ANCOVA results support hypothesis 3, meaning that learners' proficiency or existing knowledge aids them in learning better.

The individual item analysis provided detailed scoring of the participants' performance. The scoring logic was as follows: an "O" indicated learning, assigned when a participant answered incorrectly on the pre-test but correctly on the post-test; an "=" signified existing knowledge, where participants answered correctly on both tests; and an "X" indicated no learning, applied to both Correct-Error and Error-Error response patterns. Only the "O" were considered as learning occurred. The findings show that erroneous responses are higher in the explicitly learned group. This result can be interpreted as that implicit learning is more effective than explicit learning, at least within this study and seems to support the mental effort hypothesis, which claims that the retention of an inferred word meaning will be higher than that of a given word meaning (Inferring-from-context method).

However, numerous studies have already concluded that explicit learning is more effective, as discussed in section 2. Moreover, according to the IL theory, explicitly instructed groups were expected to gain higher vocabulary retention as the explicit group received more load (glossed text) than the implicit group. From the ILT point of view, this study can be interpreted as having two separate tasks: explicit and implicit. Calculating the involvement load for these two tasks following the ILT classification, the implicit group test load is 2, and the explicit group test load is 3, as there is one more load, the glossed text. Indeed, the students' needs were moderate (1), as they were told to determine the meaning of the 31 target words. There was no searching activity for the implicit group (0) but moderate searching activity for the explicit group (1), whose members were given the glossed text. Evaluation was moderate for both groups as participants were asked to undertake a gap-fill activity at post-test (1) but not a more rigorous activity like the one presented in Hulstijn and Laufer (2001). Then, why did the explicit group not outperform the implicit group? There are several guesses for the result: either the implicit group learners were good at guessing, or the explicit learning was not sufficient for the learners. Or, we can assume that learning may be too overwhelming for the explicit group, meaning the load was more than they could handle, given their level.

Looking into the number of retention, the rate of learning equates to five to six words (approximately 17% out of 31 words). Considering that one of the participants of the incidental learning study of Cho and Krashen (1994) read at least 56,000 words and gained 37 ultimately, the retention rate is not too bad. However, Cho and Krashen verified learners' gains by testing them orally and asking them to define the target words correctly. On the

contrary, the current study attempted to measure productive knowledge using a multiple-choice cloze test. While producing oral presentations and choosing correct words for blanks are aimed at measuring productive knowledge of vocabulary, having three distractors to eliminate incorrect options and producing the correct meaning orally are different levels of cognitive engagement. Therefore, we cannot decisively declare that learning occurred among participants. It could have been mere memorisation of participants, a partial learning of vocabulary, or the complete learning of the definitions provided in the text. Hence, in this study, instead of using 'learning', it may be better to use 'retention'.

The current study's findings did not comply with previous studies' results, demonstrating that implicit and explicit learning conditions did not significantly contribute to word learning. However, the study reaffirmed that language proficiency and vocabulary learning are positively related.

## **5.2. Limitations**

This section outlines the limitations of the study. It begins by examining whether the experimental environment accurately replicated a typical classroom setting. It then explores how the lack of natural reading conditions, such as the absence of dictionary use and the shift to online reading, may have hindered learners' performance. The third limitation addresses the role of repetition in vocabulary learning, seeking to understand why no significant learning occurred through reading. Following this, the adequacy of the explicit tasks is questioned, considering whether these tasks may have been insufficient or overly challenging for L2 teenage learners. Additionally, the section will discuss the discrepancy between the knowledge assessed in the pre-test and post-test and the validity of the target words used in the study.

### **5.2.1. The validity of the experiment environment**

Despite being conducted during classroom hours, the experiment was more similar to a laboratory-controlled experiment. Though the experiment was conducted in the formal classroom hours, the intervention tasks were far from similar to their ordinary classroom materials. For instance, the English class normally uses textbooks, additional study materials, and their teachers' instruction. However, during this study experiment, teachers' engagement was highly limited to a certain level, such as setting the test environment (i.e. time-keeping). They could not answer any questions from students regarding the tasks. Perhaps due to the artificial and unfamiliar setting, students might have been more nervous than usual, leading them to underperform. The result could have been different if the intervention condition was almost similar to their normal classroom environment. However,

as mentioned in Section 2.1, one thing worth referring to is that creating the perfect experimental condition in a classroom is difficult (Dekeyser, 2003). It implies that the experiment's context matters as much as the approach. The classroom environment discussion opens up possibilities for future studies and stresses the importance of carefully considering how to address these limitations before conducting classroom experiments.

### 5.2.2. Unnatural reading environments

Another limitation of this study would be the limited access to dictionary use and controlling the natural reading environment. In a natural reading environment, especially for foreign language learners, they may look up unfamiliar or new words. It could be that the deprivation of the dictionary has interfered with learners from learning naturally through research. Learners can pick up words even when searching for words. Moreover, the dictionary verification improves learning (Elgort et al., 2020; Hulstijn & Laufer, 2001). Chen (2012) stated that "dictionary use is a more effective strategy of vocabulary learning than contextual guessing" (p.1). Hulstijn and fellow researchers (Hulstijn et al., 1996) found that the chance of recalling the meaning of the words looked up in a dictionary during a reading task was higher than the marginal glossed text. Hence, the unnatural reading environment, which did not allow participants to access the dictionary, may have negatively affected participants' vocabulary learning during the intervention session. This carries pedagogical implications in L2 vocabulary learning and teaching that implicit learning of words via reading can be fortified by encouraging learners to look up a dictionary to enhance their word learning.

Furthermore, since the experiment was conducted online, the participants had to read lengthy texts on a laptop or tablet. This method of reading might have been unnatural for those who typically use these devices for visual aids rather than as their primary learning medium. Unfamiliarity with this format and potential visual fatigue could have also affected their attention and retention levels.

Hence, future studies can expand on the current study by addressing unnatural reading environments. For instance, researchers could set up diverse learning treatments: an implicit group with no clues, an implicit group with access to a dictionary, an explicit group with glossed text, and an explicit group with glossed text and access to a dictionary. Additionally, exploring the difference between paper and electronic glosses would be meaningful, reflecting that online-based instruction in the classroom is gradually increasing these days (Lee et al., 2016).

### 5.2.3. The lack of repetition and insufficient exposure

Unlike many other L2 vocabulary learning studies, which typically include multiple intervention sessions, this study did not incorporate such a design. As a result, the findings may reflect an overall decrease in learning due to the lack of repetition of words. The study involved only one session, with an immediate post-test administered just minutes after the reading task. This raises concerns that students might have memorised the words temporarily rather than genuinely learning or internalising them. Consequently, the assessment may indicate the students' short-term recall or recognition abilities rather than true vocabulary acquisition.

Hulstijn and Laufer (2001) have demonstrated that low retention rates are common in incidental vocabulary learning because a single exposure rarely results in lasting memory formation. Repeated exposure is essential for enhancing long-term recall and retrieval of new words. This is supported by research from Joseph & Nation (2018) and Waring & Takaki (2003), who agree that repeated encounters with vocabulary significantly improve acquisition by reinforcing memory. Additionally, studies on context-based vocabulary learning have shown that frequent exposure strengthens learning and helps learners understand how words are commonly used together (Bamford & Day, 1998; Laufer & Ravenhorst-Kalovski, 2010; Schmitt, 2008). Schmitt (2010) and Webb (2007) observed that a single exposure, without repetition, has minimal impact on vocabulary retention. Specifically, Webb's (2007) repetition study on vocabulary knowledge with 121 Japanese students found that at least three repetitions distinguish greatly from a single exposure. Only after seven exposures were notable significant developments in all aspects, such as orthographic form, grammar, syntax, and association. This highlights the difficulty of conducting research with authentic reading materials, which might not contain three to seven exposures of a single word, especially when adopted in a short exposure experience. Therefore, the brief, one-time exposure in this study was likely insufficient to convert recognition of target words into long-term memory. In turn, this implies that more explicit instruction might be needed. (Meganathan et al., 2019).

What is more perplexing here is that the post-test scores of participants in the implicit group were more accurate than those in the explicit group. However, the outperforming of the implicit group cannot be explained by the lack of repetition of target words. Given that the mean LexTALE scores were very similar between both groups and hypothetically considering that repeated exposure was expected to favour the explicit group due to the glossed text drawing learners' attention to the target words, it is surprising that the implicit group outperformed them. If the problem was repetition, the implicit group should have recorded less accurate scores in the post-test because the implicit group only encountered

the words. A plausible explanation could be that the 'random guessing' of the implicit group participants was more fortunate than that of the explicit group, especially among those who did not pay sufficient attention to the experiment or were not fully engaged in activating their cognitive abilities to solve the tasks.

#### 5.2.4. The adequacy of the explicit stimuli

The study's explicit learning condition may have been insufficient or too difficult. Unlike many other empirical studies that provide more than the meanings of the target words in explicit learning conditions, such as practice tasks or practice sessions are necessary, the current study only provided the marginal glossed text. Furthermore, studies that used the marginal glosses provided L1 translations in the margin for the target words. This study provided the L2 definitions. We can assume that learners may be pressured by reading L2 definitions for unfamiliar or new words, which may create a more challenging learning environment that is not distinguishable from the implicit learning environment.

In Hulstijn and Laufer's (2001) study, the explicit learning condition was very salient, for participants were involved in additional word-focused activities such as writing or fill-in-the-blank questions. Those additional tasks enhanced participants' learning after reading. Meanwhile, in this study, only the glossed text was provided to the participants, with no further activities to distinguish the implicit and explicit conditions.

Moreover, the density of L2 definitions might overwhelm the learners and eventually prevent them from attending to the target words. Jennings (2019) expressed the same concern that the definitions must be child-friendly so children do not misunderstand the target words. For future studies, it is essential to use 'more accessible, child-friendly language and teaching strategies to improve their vocabulary knowledge' (Beck, McKeown and Kucan, 2013, cited in Brownfoot, 2020).

Additionally, the study could have benefited from selecting a more engaging activity that might better capture students' attention. Incorporating more explicit instructional conditions, such as visualisation, might have led to different outcomes for the explicit learning group. For instance, in Lao and Krashen's (2000) study, participants were shown films to enhance their understanding of the texts. *The Giver*, which was adapted into a film in 2014 featuring well-known actors and celebrities like Meryl Streep, Taylor Swift, and Alexander Skarsgard, could have been a powerful additional stimulus to support learning. Additionally, incorporating other explicit conditions, such as adopting L1 translations for glossed texts, could have produced different results, aligning more closely with the

hypotheses proposed in the research questions.

### 5.2.5 Receptive and Productive Knowledge

The pre-test assessed receptive vocabulary, but the post-tests assessed productive vocabulary. Knowing a word receptively doesn't entail knowing it productively, as the two domains involve different skills. The definitions provided in the glossed text focused on the meaning of words, while the post-test format, the cloze test, concentrated on retrieving form and meaning simultaneously. Furthermore, learners had to read the sentence and decide the word more suited for the context. Therefore, the cognitive abilities engaged in the two tests are fundamentally different, making converting receptive knowledge to productive knowledge challenging for L2 learners. Additionally, one of the effects of explicit instruction was found to be a facilitation of in-depth vocabulary knowledge, meaning that learners are more prone to learn productive vocabulary when learning is situated in the explicit learning environment. (Damhuis et al., 2014)

### 5.2.6 Target words

#### Semantic ambiguity: non-literal meanings of eight words

The inclusion of eight contextually adapted words (asking for each different meaning in the pre-test and the post-test) may have been ambiguous and confusing to the participants. These "semantically ambiguous words", indicating words with more than one meaning (Eddington & Tokowicz, 2015, p. 13), and more challenging to process than the single form-meaning mapping, also called polysemous words. According to Hoffman and colleagues (2013), language use is a very complicated process; hence the versatile use of vocabulary and the infinite combinations of words often make interpreting ambiguous words extremely difficult. The authors state, "In the case of polysemous words, the ambiguity is more graded and subtle." (p.1) Moreover, the research also suggested that semantic uncertainty can greatly affect the processing and understanding of language, particularly when words are used in specific situations such as tasks that require evaluating the connection between words or grouping them. Individuals must concentrate on a specific interpretation of each word in this case. Therefore, words with multiple meanings experience challenges in processing.

The individual item-level analysis revealed that learners faced even greater challenges when word classes changed during the experiment (e.g., Noun to Verb, Verb to Noun). Lee and Federmeier's (2009) empirical study explains why the word class change brought more challenges to learners. The study explored the semantic and syntactic roles in resolving ambiguity in noun/noun (NN) homographs and noun/verb (NV)

homographs—words with more than one meaning despite sharing the same spelling. Research on NN-homographs has primarily focused on how semantic context, such as surrounding words or the overall meaning of a sentence, guides the selection of the correct meaning of an ambiguous word. In contrast, resolving ambiguity in NV-homographs depends heavily on syntactic context—specifically, whether the word functions as a noun or a verb. This syntactic ambiguity adds a significant layer of complexity, as it involves understanding the word's meaning and grammatical function. The key difference between NN-homographs and NV-homographs lies in their syntactic roles. NN-homographs present ambiguity within the same grammatical category (e.g., two different noun meanings), whereas NV-homographs involve ambiguity that crosses grammatical categories (e.g., the distinction between a noun and a verb). Therefore, while NN-homographs involve ambiguity within a single syntactic category, NV-homographs require navigating across different syntactic categories, making them inherently more complex. This complexity significantly increases the cognitive load on learners, making vocabulary acquisition in such environments particularly challenging. The need to simultaneously resolve both semantic and syntactic ambiguities suggests that learning vocabulary in the context of NV-homographs is a more demanding and intricate process, hence explaining why the learning of two words, 'pepper' and 'treat', was especially difficult for the participants.

#### The overwhelming stimuli and Participants' pre-knowledge

In addition to the semantic ambiguity, the large number of stimuli may have worn out students. Learning could have happened if half the items were repeated twice.

The higher accuracy rate of the pre-test and the lower accuracy rate of the post-test may reflect that participants knew the target vocabulary before the experiment, but possibly partially. Also, there is a possibility that participants might have learned the words during the pre-test. The multiple-choice questions could have guided learners to infer the meaning of the target words (Hulstijn, 1992). Multiple-choice questions have been empirically shown to aid L2 learners in inferring word meanings. The multiple-choice question gives learners hints, which decreases the likelihood of learners inferring incorrect meanings. Given that a cue is provided to learners or readers, it may also help those with less well-developed problem-solving skills (Hulstijn, 1992). Based on the empirical study that the Multiple Choice method is helpful for L2 learners to pick up the meaning of target words, we can also naturally infer that there was incidental learning while the participants of this study were taking the Multiple Choice pre-test. However, this does not help us understand why the implicit group learners yielded higher post-test scores than the explicit group, taking us back to limitation 5.2.3.

Back to partial knowledge, compared to knowing the simple aspect of vocabulary, deciding in what context the vocabulary should be used is much more challenging for learners, especially when they were only instructed for 15 minutes. Also, as Schmitt (2014) described in the size and depth of vocabulary knowledge study, the discrepancy between the aspects of vocabulary knowledge disappears when the frequency of the words is lower. Additionally, the demonstration of partial knowledge can be considered as the development of one or another dimension of knowledge (Bruton, 2009), which means that even though some learners failed to get correct answers in the post-test, other dimensions of knowledge may have enhanced, such as the grammar functions or idiomatic use of vocabulary.

In this context, as the two tests (pre- and post-tests) measured different aspects of vocabulary knowledge, we cannot definitively conclude that no learning has occurred. The researcher might have lost the opportunity to capture the growth of vocabulary knowledge, even if that improvement is nothing but a subtle sense of learners' which only themselves notice but cannot be described or presented in proof of learning. It means, for example, there is a possibility that the participant knew the literal meaning of the word but wasn't aware of the contextual adopted meaning in the reading text. They could have learned a simple word meaning via reading in this experiment but not the contextually adapted, semantically ambiguous meaning.

## 6. Conclusion

This study explored the effectiveness of two interventions—implicit and explicit treatments—on the vocabulary learning of Korean EFL teenage learners. Although the study initially hypothesised that vocabulary learning through reading, particularly under explicit learning conditions, would yield better retention, the findings deviated from these expectations. The statistical analysis revealed significant correlations between pre-existing knowledge, English proficiency, and the retention of target words, indicating that higher-proficiency learners tended to remember more words. However, no significant differences were observed between the implicit and explicit learning conditions. Interestingly, a significant difference was found between the pre-test and post-test scores. Still, this difference was negative, meaning that participants scored higher on the pre-test, with lower retention or learning scores on the post-test.

Both learning groups showed lower post-test scores than their pre-test scores, with the implicit group unexpectedly outperforming the explicit group in retention. Several factors could explain these unexpected results, including the mismatch of knowledge assessed in the pre-test and post-test, the overwhelming vocabulary load presented within a short treatment time (31 words within 15 minutes), the inclusion of semantically ambiguous words, and insufficient exposure to the target words due to a lack of repetition. Additionally, learners' pre-existing knowledge of the target words (an average of nine words) and the unfamiliar classroom experiment setting and unnatural reading environment may have contributed to the contradictory outcomes. There could also be other unidentified factors influencing these results.

However, one consistent finding aligned with previous studies is vocabulary learning is extremely challenging for L2 learners when limited to a single reading task. As Beck and McKeown (2007) noted, "just because it demonstrates that word learning does not occur easily," vocabulary acquisition involves multiple dimensions, from understanding the literal meaning of words to applying them appropriately in context.

These findings can inform future studies exploring the differences between implicit and explicit learning conditions in vocabulary acquisition. Should this study be replicated, conducting it within a longitudinal framework would be beneficial, where interventions are integrated into the classroom setting, providing students with sufficient and repeated exposure to target words to foster learning and have learners' receptive knowledge become productive knowledge. Additionally, incorporating qualitative interviews could offer valuable insights into the learners' cognitive processes, presenting an intriguing direction for future

research. Furthermore, using eye-tracking technology in similar future studies could serve as an additional research tool to explore participants' behaviours and whether they focus on the target words.

Lastly, the implications of this study can be applied to educational practices, curriculum development and possibly the field of L2 vocabulary learning. The findings may guide teachers in selecting the most effective instructional methods for vocabulary acquisition, potentially leading to adjustments in curriculum design that better support learners at varying proficiency levels. Furthermore, the study could influence vocabulary knowledge assessment strategies by highlighting the importance of addressing different conditions of vocabulary learning. Additionally, the study's outcomes may contribute to the empirical aspect of vocabulary acquisition, informing both existing and emerging language vocabulary learning studies by referring to limitations.

## References

- Bae, J. (2012). *Developing General Literacy Ability and Intercultural Sensitivity through English Literacy Instruction: Using Global Literature for Korean EFL Learners* [Ph.D.]. <https://www.proquest.com/docview/926193322/abstract/A01761533584D35PQ/1>
- Bamford, J., & Day, R. R. (1998). Teaching Reading. *Annual Review of Applied Linguistics*, 18, 124–141. <https://doi.org/10.1017/S0267190500003512>
- Beck, I. L., & McKeown, M. G. (2007). Increasing Young Low-Income Children's Oral Vocabulary Repertoires through Rich and Focused Instruction. *The Elementary School Journal*, 107(3), 251–271. <https://doi.org/10.1086/511706>
- Beglar, D., & Nation, P. (2007). *A vocabulary size test*.
- Bensoussan, M., & Laufer, B. (1984). Lexical Guessing in Context in EFL Reading Comprehension. *Journal of Research in Reading*, 7(1), 15–32. <https://doi.org/10.1111/j.1467-9817.1984.tb00252.x>
- Bruton, A. (2009). The Vocabulary Knowledge Scale: A Critical Analysis. *Language Assessment Quarterly*, 6(4), 288–297. <https://doi.org/10.1080/15434300902801909>
- Brybaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, 109, 104047. <https://doi.org/10.1016/j.jml.2019.104047>
- Chen, Y. (2012). Dictionary Use and Vocabulary Learning in the Context of Reading. *International Journal of Lexicography*, 25(2), 216–247. <https://doi.org/10.1093/ijl/ecr031>
- Cho, K.-S., & Krashen, S. D. (1994). Acquisition of Vocabulary from the Sweet Valley Kids Series: Adult ESL Acquisition. *Journal of Reading*, 37(8), 662–667.
- Choi, S., Kim, J., & Ryu, K. (2014). Effects of context on implicit and explicit lexical knowledge: An event-related potential study. *Neuropsychologia*, 63, 226–234. <https://doi.org/10.1016/j.neuropsychologia.2014.09.003>
- Choi, Y. (2023). Consistency of 2015 revised national curriculum, textbooks, and CSAT

- according to CEFR levels: Focusing on English reading passages. *English Teaching*, 78(1), 271–291. <https://doi.org/10.15858/engtea.78.1.202303.271>
- Chon, Y., & Shin, D. (2012). Learning Strategies and Vocabulary Knowledge of EFL Sixth and Ninth Grade Learners. *English Teaching*, 67(2), 221–246. <https://doi.org/10.15858/engtea.67.2.201207.221>
- Damhuis, C. M. P., Segers, E., & Verhoeven, L. (2014). Sustainability of Breadth and Depth of Vocabulary after Implicit versus Explicit Instruction in Kindergarten. *International Journal of Disability, Development and Education*, 61(3), 194–211. <https://doi.org/10.1080/1034912X.2014.932562>
- Dechert, H.-W., & Raupach, M. (1989). *Interlingual Processes*. Gunter Narr Verlag.
- DeKeyser, R. (2003). Implicit and Explicit Learning. In *The Handbook of Second Language Acquisition* (pp. 312–348). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470756492.ch11>
- Elgort, I., Beliaeva, N., & Boers, F. (2020). CONTEXTUAL WORD LEARNING IN THE FIRST AND SECOND LANGUAGE: DEFINITION PLACEMENT AND INFERENCE ERROR EFFECTS ON DECLARATIVE AND NONDECLARATIVE KNOWLEDGE. *Studies in Second Language Acquisition*, 42(1), 7–32. <https://doi.org/10.1017/S02722263119000561>
- Ellis, R. (2015). *Understanding second language acquisition* (Second edition.). Oxford University Press.
- Ellis, R., Loewen, S., Elder, C., Reinders, H., Erlam, R., & Philp, J. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Multilingual Matters. <https://doi.org/10.21832/9781847691767>
- Gentner, D., & Beranek, B. (1982). *Why nouns are learned before verbs: Linguistic relativity versus natural partitioning /*.
- Grabe, W., & Stoller, F. L. (2019). *Teaching and Researching Reading*. Taylor & Francis Group. <http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=5904709>

- Gyllstad, H., McLean, S., & Stewart, J. (2019, July 1). *Empirically investigating the adequacy of item sample sizes of vocabulary levels and vocabulary size tests – A bootstrapping approach: Vocab@Leuven 2019*.
- Hashimoto, B. J., & Egbert, J. (2019). More Than Frequency? Exploring Predictors of Word Difficulty for Second Language Learners. *Language Learning*, 69(4), 839–872.  
<https://doi.org/10.1111/lang.12353>
- Hirsh-Pasek, K., Michnick Golinkoff, R., & Maguire, M. J. (2006). *Action Meets Word: How Children Learn Verbs*. Oxford University Press, Incorporated.  
<http://ebookcentral.proquest.com/lib/oxford/detail.action?docID=281404>
- Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, 45(3), 718–730. <https://doi.org/10.3758/s13428-012-0278-x>
- Huckin, T., & Coady, J. (1999). INCIDENTAL VOCABULARY ACQUISITION IN A SECOND LANGUAGE: A Review. *Studies in Second Language Acquisition*, 21(2), 181–193.  
<https://doi.org/10.1017/S0272263199002028>
- Hulstijn, J. H. (1992). Retention of Inferred and Given Word Meanings: Experiments in Incidental Vocabulary Learning. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 113–125). Palgrave Macmillan UK.  
[https://doi.org/10.1007/978-1-349-12396-4\\_11](https://doi.org/10.1007/978-1-349-12396-4_11)
- Hulstijn, J. H. (2003). Incidental and Intentional Learning. In *The Handbook of Second Language Acquisition* (pp. 349–381). John Wiley & Sons, Ltd.  
<https://doi.org/10.1002/9780470756492.ch12>
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental Vocabulary Learning by Advanced Foreign Language Students: The Influence of Marginal Glosses, Dictionary Use, and Reoccurrence of Unknown Words. *The Modern Language Journal*, 80(3), 327–339. <https://doi.org/10.1111/j.1540-4781.1996.tb01614.x>
- Hulstijn, J. H., & Laufer, B. (2001). Some Empirical Evidence for the Involvement Load

- Hypothesis in Vocabulary Acquisition. *Language Learning*, 51(3), 539–558.  
<https://doi.org/10.1111/0023-8333.00164>
- Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *PSICOLOGICA*, 35(1), Article 1.
- Jeon, I. (2007). The Relationship between Korean EFL Learners' Vocabulary Ability and Vocabulary Learning Strategies. *English Teaching*, 62(1), 31–54.  
<https://doi.org/10.15858/engtea.62.1.200703.31>
- Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology*, 166, 190–211. <https://doi.org/10.1016/j.jecp.2017.08.010>
- 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.  
<https://doi.org/10.1177/1362168808089922>
- Kim, T. K., & Park, J. H. (2019). More about the basic assumptions of t-test: Normality and sample size. *Korean Journal of Anesthesiology*, 72(4), 331–335.  
<https://doi.org/10.4097/kja.d.18.00292>
- Klingner, J. K., & Vaughn, S. (2004). Strategies for struggling second-language readers. In *Adolescent literacy research and practice* (pp. 183–209).
- Ko, M. H. (1995). *Glossing in Incidental and Intentional Learning of Foreign Language Vocabulary and Reading Comprehension* [M.A.].  
<https://www.proquest.com/docview/2481234535/citation/465E9387A03D4D66PQ/1>
- Krashen, S. (1989). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language Journal*, 73(4), 440–464.  
<https://doi.org/10.2307/326879>
- Krashen, S. (2013). Reading and Vocabulary Acquisition: Supporting Evidence and Some Objections. *Iranian Journal of Language Teaching Research*, 1(1), 27–43.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*.

- Kweon, S.-O., & Kim, H.-R. (2008). *Beyond raw frequency: Incidental vocabulary acquisition in extensive reading*.
- Lao, C. Y., & Krashen, S. (2000). The impact of popular literature study on literacy development in EFL: More evidence for the power of reading. *System*, 28(2), 261–270. [https://doi.org/10.1016/S0346-251X\(00\)00011-7](https://doi.org/10.1016/S0346-251X(00)00011-7)
- Laufer, B. (2003). Vocabulary Acquisition in a Second Language: Do Learners Really Acquire Most Vocabulary by Reading? Some Empirical Evidence. *The Canadian Modern Language Review*, 59(4), 567–587. <https://doi.org/10.3138/cmlr.59.4.567>
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. <https://doi.org/10.1093/applin/22.1.1>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). *Lexical threshold revisited: Lexical text coverage, learners' vocabulary size*.
- Lee, C., & Federmeier, K. D. (2009). Wave-ering: An ERP study of syntactic and semantic context effects on ambiguity resolution for noun/verb homographs. *Journal of Memory and Language*, 61(4), 538–555. <https://doi.org/10.1016/j.jml.2009.08.003>
- Lee, H., Lee, H., & Lee, J. H. (2016). Evaluation of Electronic and Paper Textual Glosses on Second Language Vocabulary Learning and Reading Comprehension. *The Asia-Pacific Education Researcher*, 25(4), 499–507. <https://doi.org/10.1007/s40299-015-0270-1>
- Lee, H., & Lee, J. H. (2015). The Effects of Electronic Glossing Types on Foreign Language Vocabulary Learning: Different Types of Format and Glossary Information. *The Asia-Pacific Education Researcher*, 24(4), 591–601. <https://doi.org/10.1007/s40299-014-0204-3>
- Lee, M., & Choi, S. (2020). Effects of Implicit and Explicit Instruction on Morphological Awareness and Knowledge. *English Education*, 25. <https://doi.org/10.22275/SEE.25.1.01>

- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 25(1), 91–102. [https://doi.org/10.1016/S0346-251X\(96\)00063-2](https://doi.org/10.1016/S0346-251X(96)00063-2)
- Meganathan, P., Thai Yap, N., Paramasivam, S., & Jalaluddin, I. (2019). Incidental and Intentional Learning of Vocabulary among Young ESL Learners. *3L The Southeast Asian Journal of English Language Studies*, 25(4), 51–67. <https://doi.org/10.17576/3L-2019-2504-04>
- Min, H.-T. (2008). EFL Vocabulary Acquisition and Retention: Reading Plus Vocabulary Enhancement Activities and Narrow Reading. *Language Learning*, 58(1), 73–115. <https://doi.org/10.1111/j.1467-9922.2007.00435.x>
- MIRZAI, M. (2012). *Implicit Vs. Explicit Vocabulary Learning: Which Approach Serves Long-term Recall Better? 18.*
- Murphy, V. A. (2014). *Second Language Learning in the Early School Years: Trends and Contexts*. Oxford University Press.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning Word Meanings from Context during Normal Reading. *American Educational Research Journal*, 24(2), 237–270. <https://doi.org/10.2307/1162893>
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning Words from Context. *Reading Research Quarterly*, 20(2), 233–253. <https://doi.org/10.2307/747758>
- Nation, I. S. P. (2022). *Learning Vocabulary in Another Language* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781009093873>
- Nilsen, A. P., & Donelson, K. L. (2009). *Literature for today's young adults* (8th ed). Pearson.
- Paribakht, T. S., & Wesche, M. (1996). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (1st ed., pp. 174–200). Cambridge

- University Press. <https://doi.org/10.1017/CBO9781139524643.013>
- Park, A. Y., Isaacs, T., & Woodfield, H. (2018). A comparison of the effects of extensive and intensive reading approaches on the vocabulary development of Korean secondary EFL learners. *Applied Linguistics Review*, 9(1), 113–134.  
<https://doi.org/10.1515/applirev-2017-0025>
- Park, J.-E. (2001). *Korean EFL Learners' Vocabulary Learning Strategies*.
- Pigada, M., & Schmitt, N. (2006). Vocabulary Acquisition from Extensive Reading: A Case Study. *Reading in a Foreign Language*, 18(1), 1–28.
- Puig-Mayenco, E., Chaouch-Orozco, A., Liu, H., & Martín-Villena, F. (2023). The LexTALE as a measure of L2 global proficiency: A cautionary tale based on a partial replication of Lemhöfer and Broersma (2012). *Linguistic Approaches to Bilingualism*, 13(3), 299–314. <https://doi.org/10.1075/lab.22048.pui>
- Qi, S., Teng, M. F., & Fu, A. (2024). LexCH: A quick and reliable receptive vocabulary size test for Chinese Learners. *Applied Linguistics Review*, 15(2), 643–670.  
<https://doi.org/10.1515/applirev-2022-0006>
- Qian, D. D., & Lin, L. H. F. (2019). The Relationship Between Vocabulary Knowledge and Language Proficiency. In *The Routledge Handbook of Vocabulary Studies*. Routledge.
- Razali, N. M., & Wah, Y. B. (2011). *Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests*.
- Royston, P. (1992). Approximating the Shapiro-Wilk W-test for non-normality. *Statistics and Computing*, 2(3), 117–119. <https://doi.org/10.1007/BF01891203>
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (1st ed., pp. 3–32). Cambridge University Press.  
<https://doi.org/10.1017/CBO9781139524780.003>
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363.

<https://doi.org/10.1177/1362168808089921>

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.

<https://doi.org/10.1017/S0261444812000018>

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in Language Teaching* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781108569057>

Stoeckel, T., McLean, S., & Nation, P. (2021). LIMITATIONS OF SIZE AND LEVELS TESTS OF WRITTEN RECEPTIVE VOCABULARY KNOWLEDGE. *Studies in Second Language Acquisition*, 43(1), 181–203. <https://doi.org/10.1017/S027226312000025X>

Suk, N. (2017). The Effects of Extensive Reading on Reading Comprehension, Reading Rate, and Vocabulary Acquisition. *Reading Research Quarterly*, 52(1), 73–89.

<https://doi.org/10.1002/rrq.152>

Sung, H. (2013). Task-Induced Involvement Load in Korean EFL Incidental Vocabulary Learning. *언어연구*, 29(2), 269–296.

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: An intervention study. *Reading and Writing*, 22(9), 993–1019.

<https://doi.org/10.1007/s11145-008-9141-y>

Watanabe, Y. (1997). INPUT, INTAKE, AND RETENTION: Effects of Increased Processing on Incidental Learning of Foreign Language Vocabulary. *Studies in Second Language Acquisition*, 19(3), 287–307.

Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, 28(1), 46–65. <https://doi.org/10.1093/applin/aml048>

West, M. (1953). *A General service list of English words: With semantic frequencies and a supplementary word-list for the writing of popular science and technology*.

Longmans, Green.

Xiao, F. (2015). Proficiency effect on L2 pragmatic competence. *Studies in Second Language Learning and Teaching*, V(4), 557–581.

## Appendix 1: CUREC Approval

**SOCIAL SCIENCES & HUMANITIES  
INTERDIVISIONAL RESEARCH ETHICS COMMITTEE  
DEPARTMENTAL RESEARCH ETHICS COMMITTEE**

Department of Education  
15 Norham Gardens, Oxford OX2 6PY  
[student.curec@education.ox.ac.uk](mailto:student.curec@education.ox.ac.uk); [staff.curec@education.ox.ac.uk](mailto:staff.curec@education.ox.ac.uk)



Department of Education, Social Sciences Division  
University of Oxford

29 April 2024

Dear ,

### Research ethics approval

**The influence of explicit learning through literature reading in a school environment:  
EFL Adolescents' L2 vocabulary acquisition**

**Research ethics reference: EDUC\_C1A\_24\_133**

The above application has been considered on behalf of the Education Departmental Research Ethics Committee (DREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the DREC, ethics approval has now been granted for this study.

Please note the following:

**Personal data:** It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's [guidance and legal requirements](#).

**In-person activities:** Any data collection involving in-person interactions with participants must have an up-to-date fieldwork risk assessment in place; further guidance is available from the Safety Office's [website](#).

**Amendments:** Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the [SSH IDREC webpage](#).

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to [staff.curec@education.ox.ac.uk](mailto:staff.curec@education.ox.ac.uk) / [student.curec@education.ox.ac.uk](mailto:student.curec@education.ox.ac.uk) or [ethics@socsci.ox.ac.uk](mailto:ethics@socsci.ox.ac.uk).

Yours sincerely

Xin Xu

DREC member

cc: Faidra Faitaki (as Supervisor)

## Appendix 2: Opt-out form



[The influence of explicit learning through literature reading in a school environment: EFL Adolescents' L2 vocabulary acquisition 교실 환경 내 문학 읽기를 통한 명시적 학습 효과에 관한 실험: EFL 청소년들의 L2 단어 습득]

### OPT-OUT FORM

실험 참여 비동의 양식

Ethics Approval Reference: [EDUC\_C1A\_24\_133]

If you **DO NOT** want your child to participate in the above-named research study please fill out the form below and return it to the school by 20/04/2024. 귀하의 자녀가 상기 명시된 실험에 참여하지 않음을 원하시는 경우, 양식 하단의 빈칸에 정보를 입력 후 2024/4/20 까지 학교로 제출해 주시기 바랍니다.

If we do not receive an opt-out form from you by this date, your child may be invited to take part in this study, as described in the accompanying information sheet. 해당 날짜까지 비동의 양식을 제출하지 않은 경우, 귀하의 자녀는 명시된 것과 같이 해당 실험에 참여하게 됩니다. 실험에 참여하더라도 중간에 언제든지 실험을 중단할 수 있지만, 실험이 모두 끝난 후에는 취소가 불가함을 알려 드립니다. 실험 프로그램이 종료되면 모든 자료는 익명으로 보관되어 연구진조차도 개인 정보를 추적할 수 없습니다.

I, the undersigned, hereby **DO NOT** give permission for my child to take part in the study titled [The influence of explicit learning through literature reading in a school environment: EFL Adolescents' L2 vocabulary acquisition 교실 환경 내 문학 읽기를 통한 명시적 학습 효과에 관한 실험: EFL 청소년들의 L2 단어 습득]. 나, 서명인은, 아래와 같이 내 자녀가 상기 실험에 참여하는 것을 불허합니다.

Name of child (자녀 성함): \_\_\_\_\_

Name of parent/guardian (보호자 성함): \_\_\_\_\_

Signature(서명): \_\_\_\_\_ Date(서명일자): \_\_\_\_\_

Name of researcher (연구자 성명): \_\_\_\_\_

## Appendix 3: Explicit and Implicit Texts

### Explicit (Glossed text) and Implicit (No-Clue text) examples

Jonas felt nothing unusual at first. He felt only the light touch of the old man's hands on his back. He tried to relax, to breathe evenly. The room was absolutely silent, and for a moment Jonas feared that he might disgrace himself now, on the first day of his training, by falling asleep.

Then he **shivered**. He realized that the touch of the hands felt, suddenly, cold. At the same instant, breathing in, he felt the air change, and his very breath was cold. He licked his lips, and in doing so, his tongue touched the suddenly chilled air. It was very **startling**; but he was not at all frightened, now. He was filled with energy, and he breathed again, feeling the sharp intake of **frigid** air. Now, too, he could feel cold air **swirling** around his entire body. He felt it blow against his hands where they lay at his sides, and over his back. The touch of the man's hands seemed to have disappeared. Now he became aware of an entirely new sensation: **pinpricks**? No, because they were soft and without pain. Tiny, cold, featherlike feelings **peppered** his body and face. He put out his tongue again, and caught one of the dots of cold upon it. It disappeared from his awareness instantly; but he caught another, and another. The sensation made him smile.

One part of his **consciousness** knew that he was still lying there, on the bed, in the Annex room. Yet another, separate part of his being was upright now, in a sitting position, and beneath him he could feel that he was not on the soft decorated bed covering at all, but rather seated on a flat, hard surface. His hands now held (though at the same time they were still motionless at his sides) a rough, damp rope.

And he could see, though his eyes were closed. He could see a bright, **whirling** torrent of crystals in the air around him, and he could see them gather on the backs of his hands, like cold fur. His breath was visible. Beyond, through the swirl of what he now, somehow, perceived was the thing the old man had spoken of — snow — he could look out and down a great distance. He was up high someplace. The ground was thick with the furry snow, but he sat slightly above it on a hard, flat object.

Sled, he knew **abruptly**. He was sitting on a thing called sled. And the sled itself seemed to be **poised** at the top of a long, extended mound that rose from the very land where he was. Even as he thought the word "mound," his new consciousness told him hill. Then the sled, with Jonas himself upon it, began to move through the snowfall, and he understood instantly that now he was going downhill. No voice made an explanation. The experience explained itself to him.

- Shiver: the act of shaking slightly because you are frightened, cold, or ill
- Startle: to do something unexpected that surprises and sometimes worries a person or animal
- Frigid: (of weather conditions or the conditions in a room) extremely cold
- Swirl: to move quickly with a twisting, circular movement, or to make something do this
- Pinprick: a very small hole in something where a pin or needle has been pushed into it, or a sudden pain where a pin or needle has been pushed into your skin
- Consciousness: the state of being awake, thinking, and knowing what is happening around you
- Whirl: to (cause something to) turn around in circles
- Abruptly: in a sudden, unexpected, and sometimes unpleasant way
- Poise: calm confidence in a person's way of behaving, or a quality of grace (= moving in an attractive way) and balance in the way a person holds or moves their body

226

Next

Jonas felt nothing unusual at first. He felt only the light touch of the old man's hands on his back. He tried to relax, to breathe evenly. The room was absolutely silent, and for a moment Jonas feared that he might disgrace himself now, on the first day of his training, by falling asleep.

Then he shivered. He realized that the touch of the hands felt, suddenly, cold. At the same instant, breathing in, he felt the air change, and his very breath was cold. He licked his lips, and in doing so, his tongue touched the suddenly chilled air. It was very startling; but he was not at all frightened, now. He was filled with energy, and he breathed again, feeling the sharp intake of frigid air. Now, too, he could feel cold air swirling around his entire body. He felt it blow against his hands where they lay at his sides, and over his back. The touch of the man's hands seemed to have disappeared. Now he became aware of an entirely new sensation: pinpricks? No, because they were soft and without pain. Tiny, cold, featherlike feelings peppered his body and face. He put out his tongue again, and caught one of the dots of cold upon it. It disappeared from his awareness instantly; but he caught another, and another. The sensation made him smile.

One part of his consciousness knew that he was still lying there, on the bed, in the Annex room. Yet another, separate part of his being was upright now, in a sitting position, and beneath him he could feel that he was not on the soft decorated bed covering at all, but rather seated on a flat, hard surface. His hands now held (though at the same time they were still motionless at his sides) a rough, damp rope.

And he could see, though his eyes were closed. He could see a bright, whirling torrent of crystals in the air around him, and he could see them gather on the backs of his hands, like cold fur. His breath was visible. Beyond, through the swirl of what he now, somehow, perceived was the thing the old man had spoken of — snow — he could look out and down a great distance. He was up high someplace. The ground was thick with the furry snow, but he sat slightly above it on a hard, flat object.

Sled, he knew abruptly. He was sitting on a thing called sled. And the sled itself seemed to be poised at the top of a long, extended mound that rose from the very land where he was. Even as he thought the word "mound," his new consciousness told him hill. Then the sled, with Jonas himself upon it, began to move through the snowfall, and he understood instantly that now he was going downhill. No voice made an explanation. The experience explained itself to him.

225

Next

#### Appendix 4: Full reading text

The figure originally presented here cannot be made freely available via ORA because of copyright.

Jonas felt nothing unusual at first. He felt only the light touch of the old man's hands on his back.

He tried to relax, to breathe evenly. The room was absolutely silent, and for a moment Jonas feared that he might disgrace himself now, on the first day of his training, by falling asleep.

Then he shivered. He realized that the touch of the hands felt, suddenly, cold. At the same instant, breathing in, he felt the air change, and his very breath was cold. He licked his lips, and in doing so, his tongue touched the suddenly chilled air.

It was very startling; but he was not at all frightened, now. He was filled with energy, and he breathed again, feeling the sharp intake of frigid air. Now, too, he could feel cold air swirling around his entire body. He felt it blow against his hands where they lay at his sides, and over his back.

The touch of the man's hands seemed to have disappeared.

Now he became aware of an entirely new sensation: pinpricks? No, because they were soft and without pain. Tiny, cold, featherlike feelings peppered his body and face. He put out his tongue again, and caught one of the dots of cold upon it. It disappeared from his awareness instantly; but he caught another, and another. The sensation made him smile.

One part of his consciousness knew that he was still lying there, on the bed, in the Annex room. Yet another, separate part of his being was upright now, in a sitting position, and beneath him he could feel that he was not on the soft decorated

bedcovering at all, but rather seated on a flat, hard surface. His hands now held (though at the same time they were still motionless at his sides) a rough, damp rope.

And he could *see*, though his eyes were closed. He could see a bright, whirling torrent of crystals in the air around him, and he could see them gather on the backs of his hands, like cold fur.

His breath was visible.

Beyond, through the swirl of what he now, somehow, perceived was the thing the old man had spoken of — *snow*— he could look out and down a great distance. He was up high someplace. The ground was thick with the furry snow, but he sat slightly above it on a hard, flat object.

Sled, he knew abruptly. He was sitting on a thing called *sled*. And the sled itself seemed to be poised at the top of a long, extended mound that rose from the very land where he was. Even as he thought the word "mound," his new consciousness told him *hill*.

Then the sled, with Jonas himself upon it, began to move through the snowfall, and he understood instantly that now he was going downhill. No voice made an explanation. The experience explained itself to him.

His face cut through the frigid air as he began the descent, moving through the substance called snow on the vehicle called sled, which propelled itself on what he now knew without doubt to be *runners*.

Comprehending all of those things as he sped down-ward, he was free to enjoy the breathless glee that over-whelmed him: the speed, the clear cold air, the total silence, the feeling of balance and excitement and peace.

Then, as the angle of incline lessened, as the mound — the *hill*— flattened, nearing the bottom, the sled's for-ward motion slowed. The snow was piled now around it, and he pushed with his body, moving it forward, not wanting the exhilarating ride to end.

Finally the obstruction of the piled snow was too much for the thin runners of the sled, and he came to a stop. He sat there for a moment, panting, holding the rope in his cold hands. Tentatively he opened his eyes — not his snow-hill- sled eyes, for they had been open throughout the strange ride. He opened his ordinary eyes, and saw that he was still on the bed, that he had not moved at all.

The old man, still beside the bed, was watching him.

'How do you feel ' he asked.

Jonas sat up and tried to answer honestly. "Surprised," he said, after a moment.

The old man wiped his forehead with his sleeve. "Whew," he said. "It was exhausting. But you know, even transmitting that tiny memory to you — I think it lightened me just a

little."

"Do you mean — you did say I could ask questions:" The man nodded, encouraging his question.

"Do you mean that now you don't have the memory of it— of that ride on the sled — anymore?"

"That's right. A little weight off this old body."

"But it was such fun! And now you don't have it any- more! I *took* it from you!"

But the old man laughed. "All I gave you was one ride, on one sled, in one snow, on one hill. I have a whole world of them in my memory. I could give them to you one by one, a thousand times, and there would still be more."

"Are you saying that I — I mean we — could do it again?" Jonas asked. "I'd really like to. I think I could steer, by pulling the rope. I didn't try this time, because it was so new.

The old man, laughing, shook his head. "Maybe another day, for a treat. But there's no time, really, just to play. I only wanted to begin by showing you how it works.

"Now," he said, turning businesslike, "Lie back down. I want to — "

Jonas did. He was eager for whatever experience would come next. But he had, suddenly, so many questions.

"Why don't we have snow, and sleds, and hills." he asked. "And when did we, in the past? Did my parents have sleds when they were young? Did you?"

The old man shrugged and gave a short laugh. "No," he told Jonas. "It's a very distant memory. That's why it was so exhausting — I had to tug it forward from many generations back. It was given to me when I was a new Receiver, and the previous Receiver had to pull it through a long time period, too."

"But what happened to those things? Snow, and the rest of it?"

"Climate Control. Snow made growing food difficult, limited the agricultural periods. And unpredictable weather made transportation almost impossible at times. It wasn't a practical thing, so it became obsolete when we went to Sameness.

"And hills, too," he added. "They made conveyance of goods unwieldy. Trucks; buses. Slowed them down. So — " He waved his hand, as if a gesture had caused hills to disappear. "Sameness," he concluded.

Jonas frowned. "I wish we had those things, still. Just now and then."

The old man smiled. "So do I" he said. "But that choice is not ours."

"But sir," Jonas suggested, "since you have so much power — "

The man corrected him. "Honor," he said firmly. "I have great honor. So will you. But you will find that that is not the same as power.

"Lie quietly now. Since we've entered into the topic of climate, let me give you something else. And this time I'm not going to tell you the name of it, because I want to test the receiving. You should be able to perceive the name without being told. I gave away snow and sled and down- hill and runners by telling them to you in advance."

Without being instructed, Jonas closed his eyes again. He felt the hands on his back again. He waited.

Now it came more quickly, the feelings. This time the hands didn't become cold, but instead began to feel warm on his body. They moistened a little. The warmth spread, extending across his shoulders, up his neck, onto the side of his face. He could feel it through his clothed parts, too: a pleasant, all-over sensation; and when he licked his lips this time, the air was hot and heavy. He didn't move. There was no sled. His posture didn't change. He was simply alone someplace, out of doors, lying down, and the warmth came from far above. It was not as exciting as the ride through the snowy air; but it was pleasurable and comforting.

Suddenly he perceived the word for it: *sunshine*. He perceived that it came from the sky.

Then it ended.

"Sunshine," he said aloud, opening his eyes.

"Good. You did get the word. That makes my job easier. Not so much explaining."

"And it came from the sky."

"That's right," the old man said. "Just the way it used to. "Before Sameness.

Before Climate Control," Jonas added.

The man laughed. "You receive well, and learn quickly. I'm very pleased with you. That's enough for today, I think. We're off to a good start."

There was a question bothering Jonas. "Sir," he said, "The Chief Elder told me — she told everyone — and you told me, too, that it would be painful. So I was a little scared. But it didn't hurt at all. I really enjoyed it." He looked quizzically at the old man.

The man sighed. "I started you with memories of pleasure. My previous failure gave me the wisdom to do that." He took a few deep breaths. "Jonas," he said, "it *will* be painful. But it need not be painful yet."

"I'm brave. I really am." Jonas sat up a little straighter. The old man looked at him for

a moment. He smiled. "I can see that," he said. "Well, since you asked the question — I think I have enough energy for one more transmission.

"Lie down once more. This will be the last today."

Jonas obeyed cheerfully. He closed his eyes, waiting, and felt the hands again; then he felt the warmth again, the sunshine again, coming from the sky of this other consciousness that was so new to him. This time, as he lay basking in the wonderful warmth, he felt the passage of time. His real self was aware that it was only a minute or two; but his other, memory-receiving self felt hours pass in the sun. His skin began to sting. Restlessly he moved one arm, bending it, and felt a sharp pain in the crease of his inner arm at the elbow.

"Ouch," he said loudly, and shifted on the bed. "Owwww," he said, wincing at the shift, and even mm-ing his mouth to speak made his face hurt.

He knew there was a word, but the pain kept him from grasping it.

Then it ended. He opened his eyes, wincing with discomfort. "It hurt," he told the man, "and I couldn't get the word for it."

"It was sunburn," the old man told him.

"It hurt *a lot*," Jonas said, "but I'm glad you gave it to me. It was interesting. And now I understand better, what it meant, that there would be pain."

The man didn't respond. He sat silently for a second. Finally he said, "Get up, now. It's time for you to go home."

They both walked to the center of the room. Jonas put his tunic back on. "Goodbye, sir," he said. "Thank you for my first day."

The old man nodded to him. He looked drained, and a little sad.

"Sir?" Jonas said shyly.

"Yes? Do you have a question?"

"It's just that I don't know your name. I thought you were The Receiver, but you say that now *I'm* The Receiver. So I don't know what to call you."

The man had sat back down in the comfortable upholstered chair. He moved his shoulders around as if to ease away an aching sensation. He seemed terribly weary.

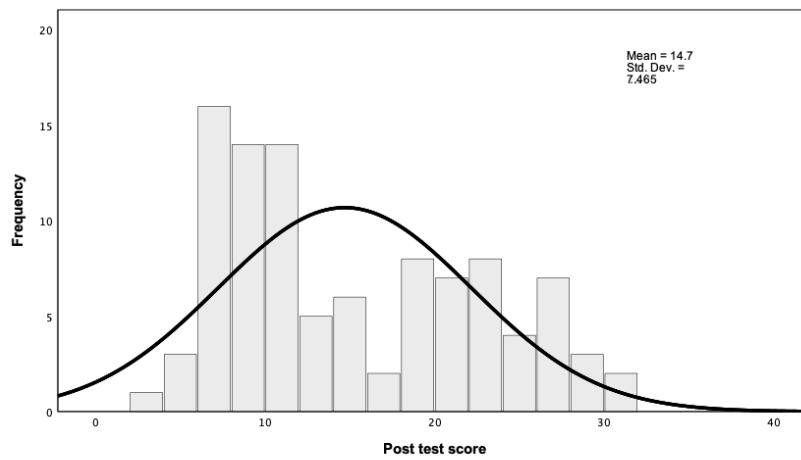
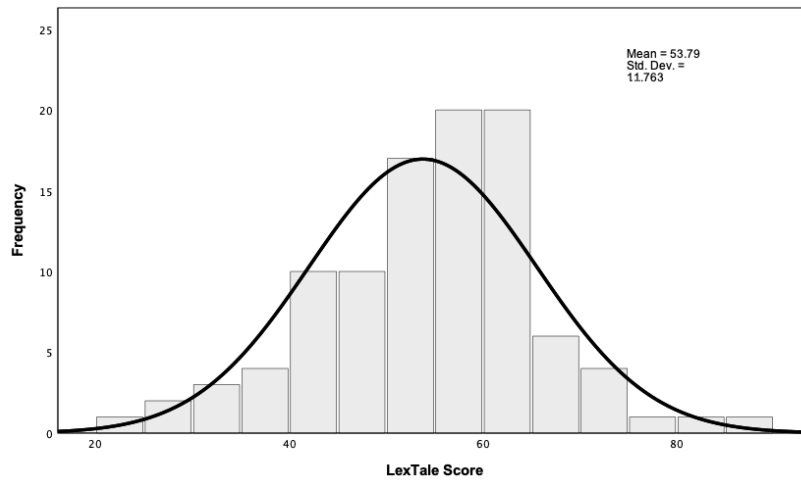
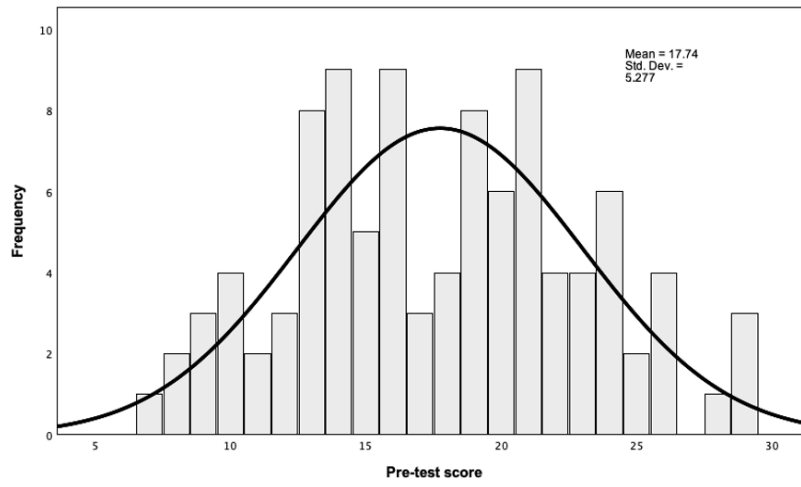
"Call me The Giver," he told Jonas.

**Appendix 5. Target words glossary (A: adjective, V: verb, N: noun)**

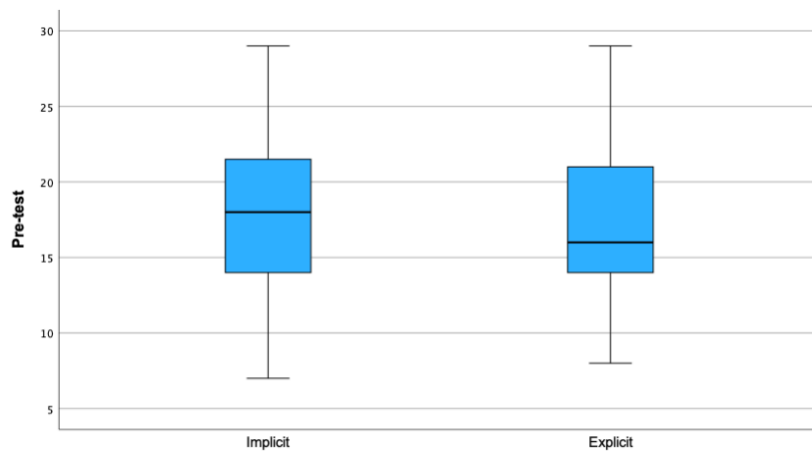
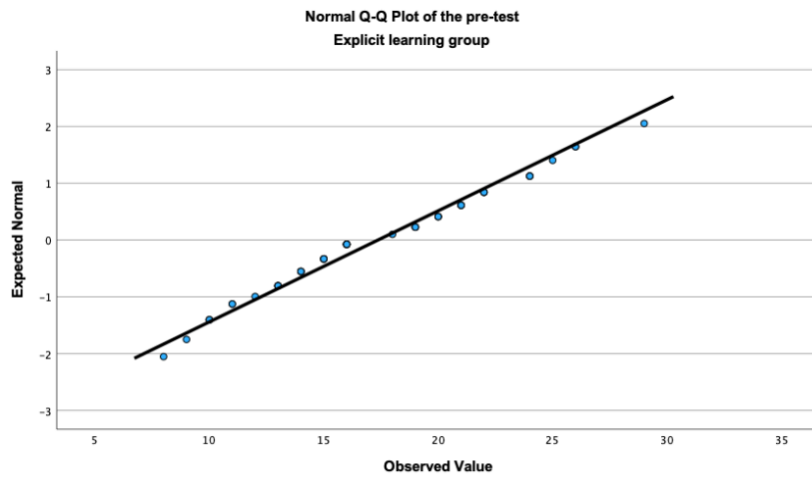
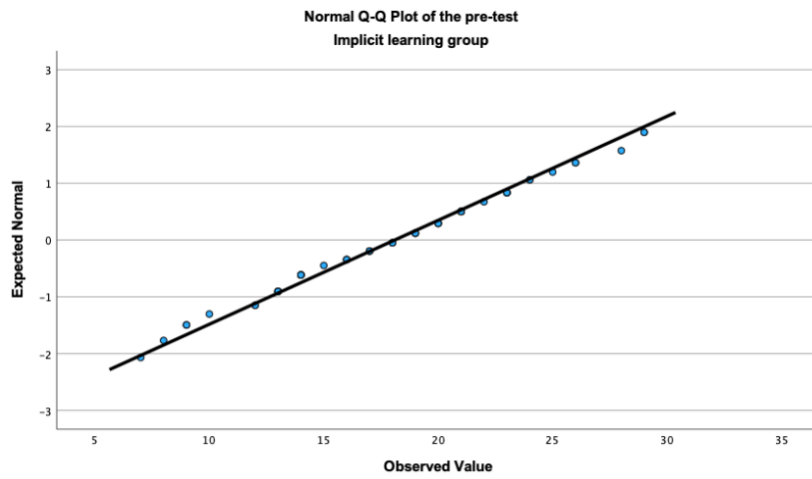
Abruptly	A
Bask	V
Comprehend	V
Consciousness	N
Conveyance	N
Crease	N
Descent	N
Drain	V
Exhilarate	V
Firmly	A
Frigid	A
Glee	N
Grasp	V
Obsolete	A
Overwhelm	V
Pant	V
Pepper	N
Perceive	V
Pinprick	N
Poise	N
Quizzically	A
Shiver	V
Shrug	V
Startle	V
Swirl	V
Tentatively	A
Treat	V
Tug	V
Unwieldy	A
Whirl	V
Wince	V

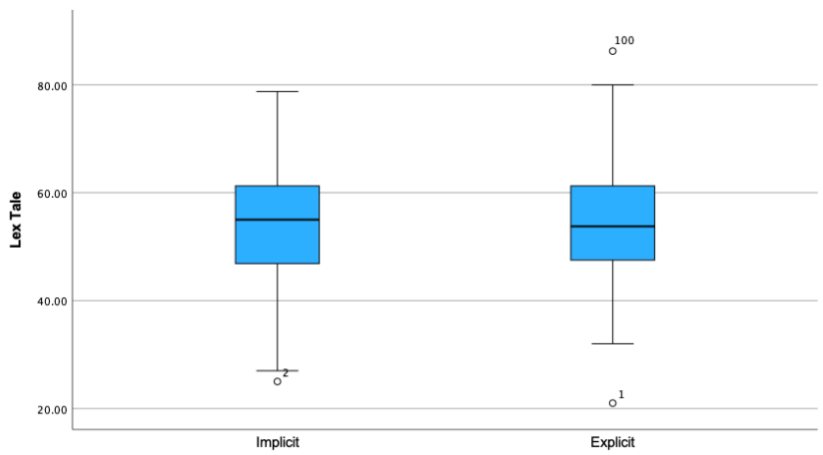
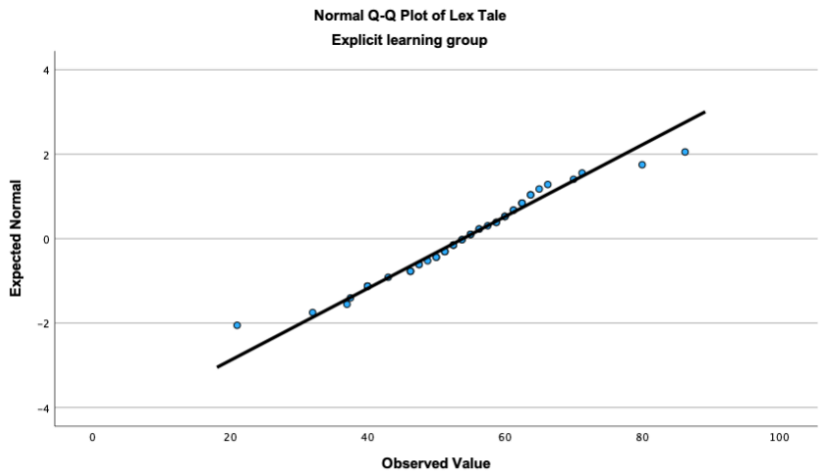
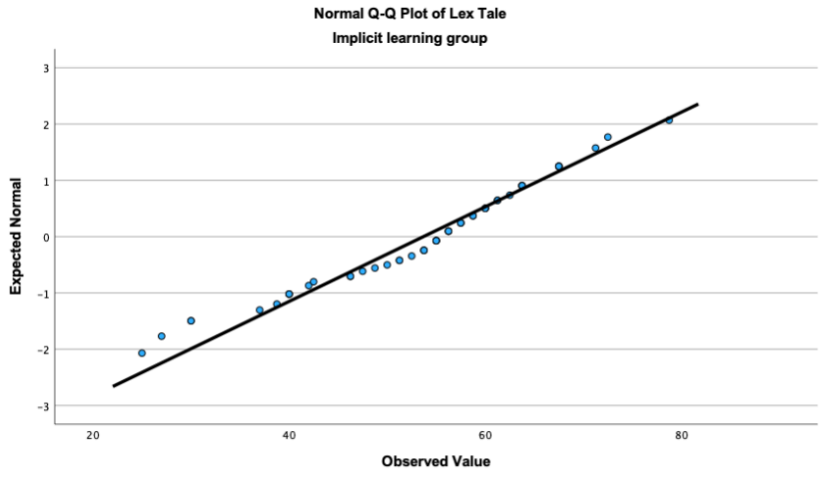
## Appendix 6. Statistics tables and charts

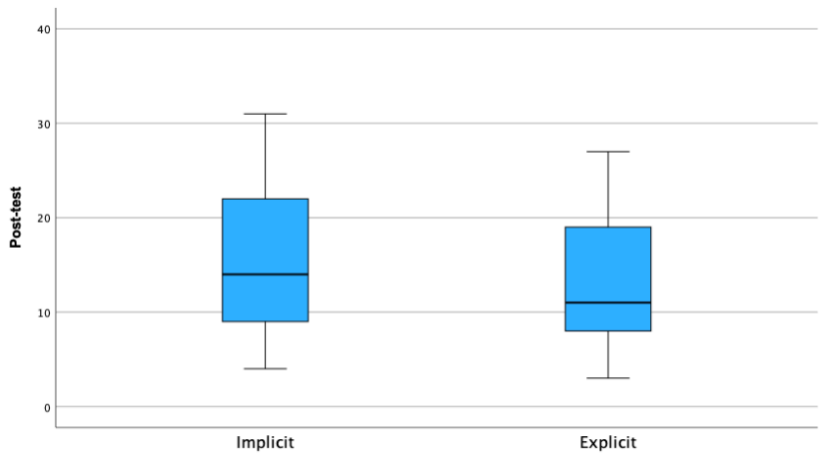
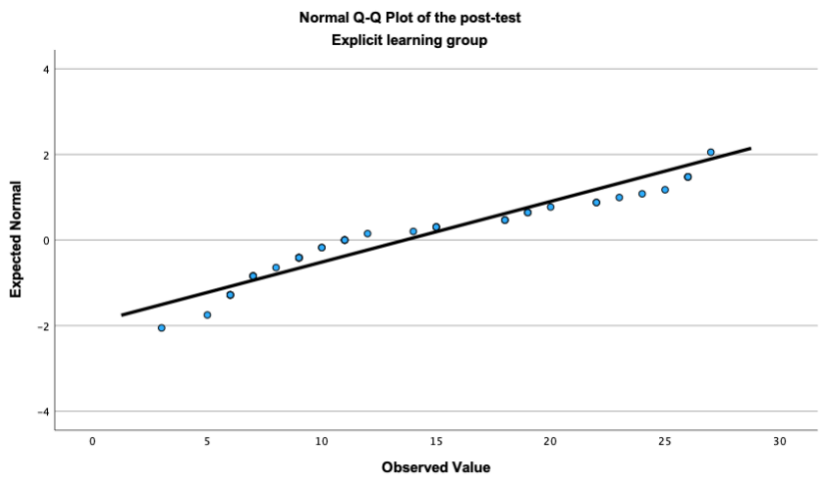
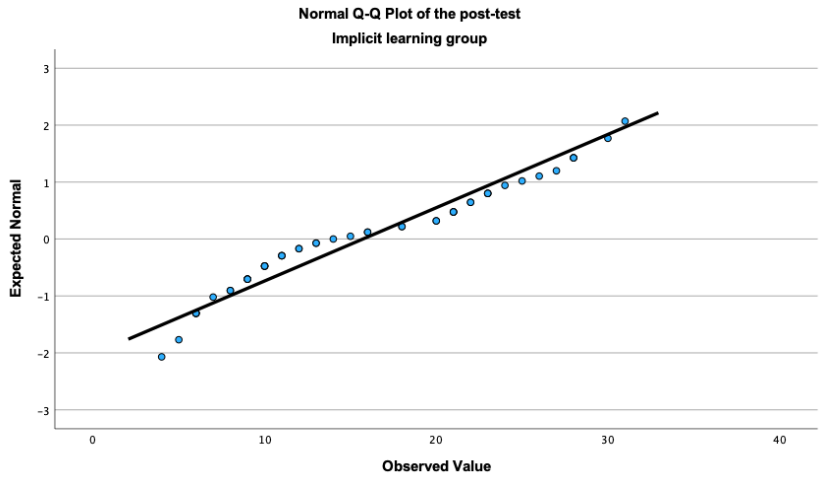
### 6.1. Histograms: test score distribution



## 6.2. Shapiro-Wilk Test: Q-Q Plots







## 6.3. Independent samples t-test results

**Group statistics**

	Group ID	N	Mean	Std. Deviation	Std. Error Mean
Pre-test	0	51	18.1	5.46	0.765
	1	49	17.37	5.11	0.73
LexTale	0	51	53.6961	11.89509	1.66565
	1	49	53.8878	11.74701	1.67814
Post-test	0	51	15.73	7.769	1.088
	1	49	13.63	7.055	1.008

## Independent Sample Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Pre-test	Equal variances assumed	.033	.857	.690	98	.246	.492	.731	1.058	-1.370	2.831
	Equal variances not assumed			.691	97.935	.246	.491	.731	1.057	-1.367	2.828
Lex Tale	Equal variances assumed	.067	.796	-.081	98	.468	.936	-.19168	2.36503	-4.88501	4.50165
	Equal variances not assumed			-.081	97.924	.468	.936	-.19168	2.36443	-4.88386	4.50051
Post-test	Equal variances assumed	1.017	.316	1.408	98	.081	.162	2.093	1.486	-.856	5.042
	Equal variances not assumed			1.411	97.695	.081	.161	2.093	1.483	-.850	5.036

### Independent Samples Effect Sizes

		Standardizer <sup>a</sup>	Point Estimate	95% Confidence Interval	
				Lower	Upper
Pre-test	Cohen's d	5.291	.138	-.255	.530
	Hedges' correction	5.332	.137	-.253	.526
	Glass's delta	5.110	.143	-.251	.535
Lex Tale	Cohen's d	11.82280	-.016	-.408	.376
	Hedges' correction	11.91425	-.016	-.405	.373
	Glass's delta	11.74701	-.016	-.408	.376
Post-test	Cohen's d	7.428	.282	-.113	.675
	Hedges' correction	7.486	.280	-.112	.670
	Glass's delta	7.055	.297	-.101	.692

a. The denominator used in estimating the effect sizes.

Cohen's d uses the pooled standard deviation.

Hedges' correction uses the pooled standard deviation, plus a correction factor.

Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

## 6.4. A 2X2 mixed ANOVA

**Descriptive statistics**

	Group ID	Mean	Std. Deviation	N
Pre-test	Implicit	18.10	5.46	51
	Explicit	17.37	5.11	49
	Total	17.74	5.28	100
Post-test	Implicit	15.73	7.769	51
	Explicit	13.63	7.06	49
	Total	14.7	7.465	100

**Multivariate Tests<sup>a</sup>**

	Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	0.165	19.363 <sup>b</sup>	1	98	<.001	0.165
	Wilks' Lambda	0.835	19.363 <sup>b</sup>	1	98	<.001	0.165
	Hotelling's Trace	0.198	19.363 <sup>b</sup>	1	98	<.001	0.165
	Roy's Largest Root	0.198	19.363 <sup>b</sup>	1	98	<.001	0.165
Time * GroupID	Pillai's Trace	0.01	.963 <sup>b</sup>	1	98	0.329	0.01
	Wilks' Lambda	0.99	.963 <sup>b</sup>	1	98	0.329	0.01
	Hotelling's Trace	0.01	.963 <sup>b</sup>	1	98	0.329	0.01
	Roy's Largest Root	0.01	.963 <sup>b</sup>	1	98	0.329	0.01

a. Design: Intercept + GroupID

Within Subjects Design: Time

b. Exact statistic

**Mauchly's Test of Sphericity<sup>a</sup>**

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>b</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Time	1	0	0	.	1	1	1

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept + GroupID

Within Subjects Design: Time

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

**Tests of Within-Subjects Effects**

	Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	Sphericity Assumed	466.044	1	466.044	19.363	<.001	0.165
	Greenhouse-Geisser	466.044	1	466.044	19.363	<.001	0.165
	Huynh-Feldt	466.044	1	466.044	19.363	<.001	0.165
	Lower-bound	466.044	1	466.044	19.363	<.001	0.165
Time * GroupID	Sphericity Assumed	23.184	1	23.184	0.963	0.329	0.01
	Greenhouse-Geisser	23.184	1	23.184	0.963	0.329	0.01
	Huynh-Feldt	23.184	1	23.184	0.963	0.329	0.01
	Lower-bound	23.184	1	23.184	0.963	0.329	0.01
Error (Time)	Sphericity Assumed	2358.736	98	24.069			
	Greenhouse-Geisser	2358.736	98	24.069			
	Huynh-Feldt	2358.736	98	24.069			
	Lower-bound	2358.736	98	24.069			

**Test of Within-Subjects Contrasts**

Source	Time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	Level 2 vs. Level 1	932.087	1	932.087	19.363	<.001	0.165
Time * GroupID	Level 2 vs. Level 1	46.367	1	46.367	0.963	0.329	0.01
Error(Time)	Level 2 vs. Level 1	4717.473	98	48.137			

**Levene's Test of Equality of Error Variances<sup>a</sup>**

		Levene Statistic	df1	df2	Sig.
Pre-test	Based on Mean	0.033	1	98	0.857
	Based on Median	0.09	1	98	0.764
	Based on Median and with adjusted df	0.09	1	97.993	0.764
	Based on trimmed mean	0.035	1	98	0.852
Post-test	Based on Mean	1.017	1	98	0.316
	Based on Median	1.212	1	98	0.274
	Based on Median and with adjusted df	1.212	1	95.794	0.274
	Based on trimmed mean	1.043	1	98	0.31

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

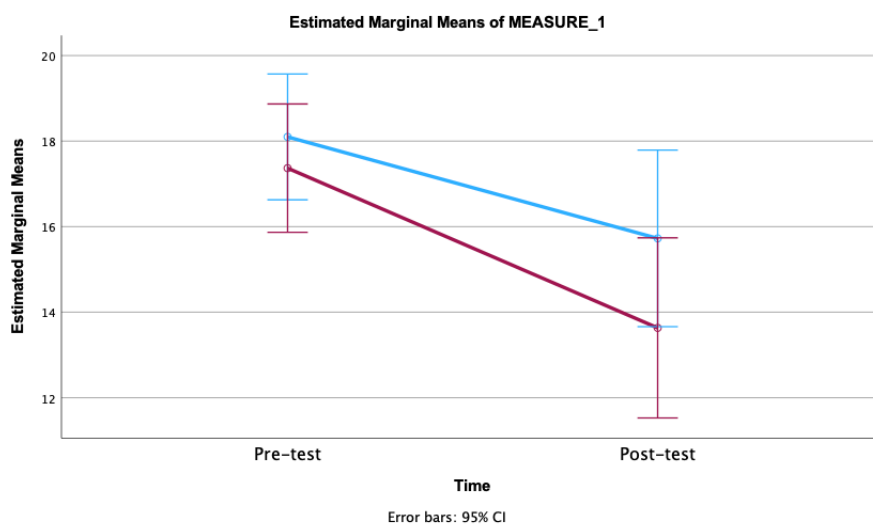
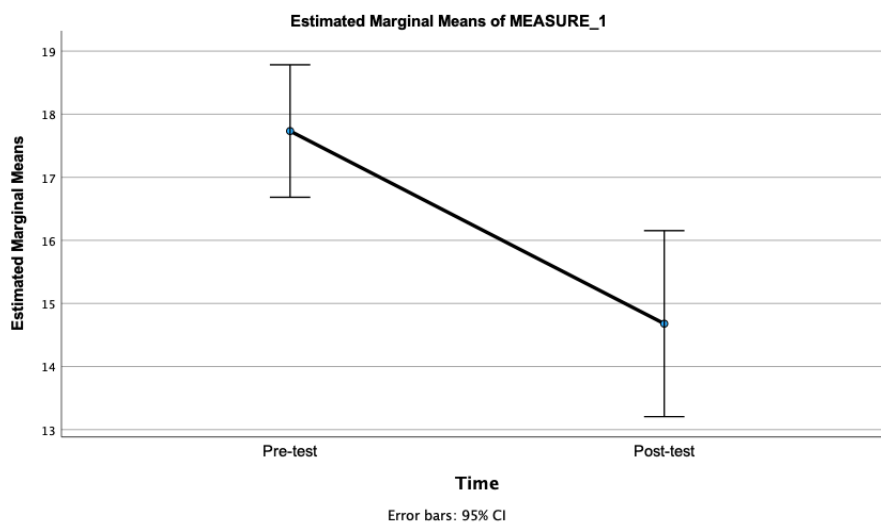
a. Design: Intercept + GroupID

Within Subjects Design: Time

## Test of Between-Subjects Effects

Transformed Variable: Average

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Intercept	26252.557	1	26252.557	888.272	<.001	0.901
GroupID	49.807	1	49.807	1.685	0.197	0.017
Error	2896.353	98	29.555			



## 6.5. ANCOVA

**Levene's Test of Equality of Error Variances<sup>a</sup>**

Dependent Variable: Post-test scores

F	<i>df1</i>	<i>df2</i>	Sig.
0.763	1	98	0.384

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Definitioncorrect + LexTale + GroupID

**Tests of Between-Subjects Effects**

Dependent Variable: Post-test scores

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1535.302 <sup>a</sup>	3	511.767	12.339	<.001	0.278
Intercept	41.845	1	41.845	1.009	0.318	0.01
Definitioncorrect	600.839	1	600.839	14.486	<.001	0.131
LexTale	351.449	1	351.449	8.474	0.004	0.081
GroupID	77.135	1	77.135	1.86	0.176	0.019
Error	3981.698	96	41.476			
Total	27126	100				
Corrected Total	5517	99				

a. R Squared = .278 (Adjusted R Squared = .256)

**Test Results**

Dependent Variable: Post-test scores

Source	Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Contrast	77.135	1	77.135	1.86	0.176	0.019
Error	3981.698	96	41.476			

**Estimates**

Dependent Variable: Post-test scores

Group ID	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
0	15.563 <sup>a</sup>	0.903	13.771	17.356
1	13.801 <sup>a</sup>	0.921	11.972	15.63

a. Covariates appearing in the model are evaluated at the following values: Pre-test = 17.74,

Lex Tale = 53.7900.

## 6.6 Individual item level analysis

**Individual item and participants analysis: pre-test and the post-test scores (Participants 1 ~ 20)**

Target words	PoS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Abruptly	A	X	=	X	X	X	X	=	X	X	X	O	=	X	O	X	X	X	O	X	X
Bask	V	X	X	X	O	=	X	=	X	X	X	X	=	=	X	O	O	X	O	X	O
Comprehend	V	X	X	=	=	X	X	=	X	=	X	X	=	X	O	O	X	=	X	O	X
Consciousness	N	=	X	=	X	=	X	=	X	=	X	X	=	X	O	=	O	X	=	=	=
Conveyance	N	O	X	X	O	=	X	=	X	X	X	X	O	=	=	=	=	O	O	X	=
<b>Crease</b>	<b>N</b>	O	X	X	X	X	X	=	X	=	X	X	O	X	O	X	X	X	O	X	X
Descent	N	X	=	=	=	X	O	=	X	X	O	=	X	X	O	X	X	X	O	X	X
<b>Drain</b>	<b>V</b>	=	X	=	=	O	X	=	X	=	X	=	=	X	=	O	X	X	O	=	O
Exhilarate	V	X	X	O	X	O	=	=	X	X	X	X	O	X	X	=	X	X	=	=	=
<b>Firmly</b>	<b>A</b>	=	X	O	O	X	X	=	X	=	X	X	=	X	O	X	=	X	O	X	O
Frigid	A	O	X	O	X	O	X	=	O	X	X	=	O	X	=	=	X	=	O	=	=
Glee	N	X	X	X	=	X	X	=	X	X	X	O	=	=	O	X	X	X	O	X	=
<b>Grasp</b>	<b>V</b>	X	=	=	X	=	X	=	X	=	=	X	=	O	=	=	=	X	=	O	X
Obsolete	A	X	X	X	=	O	X	=	X	X	X	X	O	X	X	X	X	X	O	=	O

Overwhelm	V	X	X	=	=	=	=	=	X	X	=	X	=	=	=	X	=	X	=	=	=
Pant	V	X	X	O	O	X	O	=	O	X	X	X	=	=	O	X	=	=	X	O	X
<b>Pepper</b>	<b>N</b>	X	X	X	=	X	X	=	X	X	=	X	X	X	=	X	X	X	=	=	X
Perceive	V	X	=	=	=	X	X	=	X	X	O	O	=	X	=	X	X	X	O	=	=
<b>Pinprick</b>	<b>N</b>	=	X	X	O	X	=	=	X	=	X	X	=	X	=	X	X	X	=	=	X
Poise	N	X	X	X	X	X	=	=	X	X	=	X	X	O	O	X	X	X	=	X	X
Quizzically	A	X	=	X	=	X	X	=	X	=	X	X	=	=	=	=	=	O	=	=	=
Shiver	V	=	X	O	=	X	X	=	O	=	X	X	=	X	=	=	O	X	O	=	=
Shrug	V	O	X	X	=	X	X	=	=	X	O	X	=	X	O	=	X	O	=	O	=
Startle	V	X	X	X	O	X	O	=	X	X	X	X	X	=	X	O	X	X	X	X	X
Swirl	V	X	X	X	X	=	X	=	X	X	=	X	O	X	O	X	X	X	O	=	O
Tentatively	A	X	=	X	=	X	X	O	X	X	X	X	=	X	O	O	X	X	=	=	O
<b>Treat</b>	<b>V</b>	X	=	X	=	=	X	=	X	X	X	X	X	X	O	X	X	X	O	X	X
<b>Tug</b>	<b>V</b>	X	X	X	=	X	X	=	X	=	X	=	X	O	O	X	X	X	O	=	=
Unwieldy	A	=	X	O	X	=	=	=	=	X	X	X	X	X	O	X	X	X	=	X	=
Whirl	V	X	X	X	=	X	X	=	X	X	O	X	=	X	X	=	X	X	=	=	=
Wince	V	X	=	X	X	X	=	O	X	X	X	X	O	=	=	O	X	X	O	=	X

---

---

<b>Learned</b>	4	0	6	6	4	3	2	3	0	4	3	7	3	15	6	3	3	16	4	6
<b>Not-learned</b>	21	23	18	10	19	22	0	26	21	22	24	7	20	5	16	22	25	3	11	12
<b>Prior knowledge</b>	6	8	7	15	8	6	29	2	10	5	4	17	8	11	9	6	3	12	16	13
<b>Learned %</b>	13%	0%	19%	19%	13%	10%	6%	10%	0%	13%	10%	23%	10%	48%	19%	10%	10%	52%	13%	19%
<b>Not-learned %</b>	68%	74%	58%	32%	61%	71%	0%	84%	68%	71%	77%	23%	65%	16%	52%	71%	81%	10%	35%	39%
<b>Prior-knowledge %</b>	19%	26%	23%	48%	26%	19%	94%	6%	32%	16%	13%	55%	26%	35%	29%	19%	10%	39%	52%	42%

---

## Participants 21~40

Target words	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Abruptly	X	=	O	O	X	X	X	X	X	X	X	X	X	X	X	O	X	O	X	=
Bask	=	X	=	O	O	O	X	X	X	X	O	=	X	O	X	=	O	O	X	O
Comprehend	=	=	O	=	X	=	X	=	X	O	=	=	X	X	X	=	=	=	X	=
Consciousness	=	=	O	=	=	X	O	X	X	X	=	X	=	=	X	=	O	X	X	=
Conveyance	=	X	X	O	X	=	X	X	X	X	X	O	X	O	X	=	O	=	X	=
<b>Crease</b>	X	O	O	=	X	X	X	X	=	X	O	O	X	O	X	O	O	O	X	O
Descent	=	O	O	X	O	X	X	X	X	X	X	X	X	X	X	X	=	=	X	X
<b>Drain</b>	=	=	X	O	=	=	=	X	X	X	=	O	X	X	X	=	X	O	=	=
Exhilarate	O	X	O	=	O	X	X	X	X	X	O	=	X	O	X	=	X	X	X	X
<b>Firmly</b>	=	O	O	O	X	O	X	X	O	X	O	=	X	O	X	=	O	O	X	=
Frigid	=	=	=	=	X	O	X	X	O	X	O	X	X	X	=	O	O	O	O	=
Glee	=	O	X	O	O	O	X	=	X	X	=	=	X	=	X	O	X	=	X	O
<b>Grasp</b>	=	=	=	=	X	=	X	X	O	O	=	=	X	O	X	=	X	=	O	=
Obsolete	=	X	O	O	X	=	O	X	O	X	X	=	X	O	X	=	=	=	X	=

Overwhelm	=	X	X	=	X	=	X	X	X	=	X	=	X	O	O	=	X	=	=	=
Pant	=	X	X	O	X	X	X	X	=	X	O	X	X	X	X	O	O	O	X	X
<b>Pepper</b>	X	X	=	O	X	X	X	X	O	X	=	X	X	X	X	X	X	=	X	X
Perceive	=	O	=	X	X	X	X	X	O	X	X	=	X	X	X	=	=	O	X	=
<b>Pinprick</b>	X	O	=	O	=	O	X	X	X	X	=	O	=	X	X	=	=	O	X	=
Poise	X	X	O	X	X	X	X	O	X	X	X	X	X	=	X	O	X	=	X	X
Quizzically	=	O	X	=	X	=	=	X	O	X	=	=	=	O	=	=	X	=	X	=
Shiver	=	=	X	X	X	=	X	X	X	X	=	=	X	O	X	=	X	=	O	=
Shrug	=	X	O	=	=	O	=	X	X	X	=	O	X	O	O	O	O	O	X	=
Startle	=	O	O	X	X	O	X	X	X	O	=	X	=	X	X	=	O	=	X	=
Swirl	=	O	X	X	X	X	X	X	X	X	=	X	O	X	X	X	O	X	O	
Tentatively	O	=	X	O	X	O	O	X	X	=	=	=	X	O	O	=	=	X	=	=
<b>Treat</b>	=	=	O	X	X	X	X	=	O	O	=	O	X	=	X	=	X	X	O	=
<b>Tug</b>	=	O	=	=	X	=	X	X	X	X	=	X	O	X	=	=	X	X	=	
Unwieldy	=	=	=	=	=	O	=	X	X	X	=	O	X	X	X	O	=	=	=	=
Whirl	=	O	X	X	=	X	X	O	X	X	=	X	O	X	=	=	=	O	=	
Wince	=	X	X	=	X	X	X	O	X	X	X	O	=	X	O	O	X	=	=	=

---

---

<b>Learned</b>	2	11	12	11	4	9	3	3	8	4	6	8	0	15	4	9	9	11	5	4
<b>Not-learned</b>	5	10	11	8	21	13	24	25	21	25	11	8	26	12	25	3	13	5	21	5
<b>Prior knowledge</b>	24	10	8	12	6	9	4	3	2	2	14	15	5	4	2	19	9	15	5	22
<b>Learned %</b>	6%	35%	39%	35%	13%	29%	10%	10%	26%	13%	19%	26%	0%	48%	13%	29%	29%	35%	16%	13%
<b>Not-learned %</b>	16%	32%	35%	26%	68%	42%	77%	81%	68%	81%	35%	26%	84%	39%	81%	10%	42%	16%	68%	16%
<b>Prior-knowledge %</b>	77%	32%	26%	39%	19%	29%	13%	10%	6%	6%	45%	48%	16%	13%	6%	61%	29%	48%	16%	71%

---

## Participants 41 ~ 60

Target words	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Abruptly	X	X	X	X	=	O	X	X	=	X	X	X	X	X	X	X	X	=	=	X
Bask	X	X	X	X	=	O	X	O	O	X	X	O	O	O	X	O	O	=	X	=
Comprehend	X	=	X	X	=	=	=	=	=	X	=	=	X	=	X	=	X	=	=	=
Consciousness	X	X	X	X	=	=	X	X	=	X	X	=	=	=	X	=	X	=	=	X
Conveyance	X	X	X	X	=	O	O	X	O	X	=	=	=	X	X	=	X	=	=	=
<b>Crease</b>	X	X	X	X	O	O	O	X	X	X	=	O	O	X	X	O	X	O	O	X
Descent	X	O	O	O	=	X	X	X	O	X	=	O	=	X	X	O	X	=	=	X
<b>Drain</b>	X	X	X	X	O	=	X	X	=	X	X	=	O	=	X	O	X	=	=	=
Exhilarate	X	X	X	X	O	O	X	O	=	X	=	=	=	=	X	O	O	O	=	=
<b>Firmly</b>	X	=	X	O	X	O	X	X	=	X	=	=	O	=	O	=	X	=	O	=
Frigid	=	X	X	X	O	X	O	X	=	=	=	=	O	=	X	=	X	=	X	=
Glee	X	X	X	O	=	O	X	X	=	X	=	X	=	=	X	O	X	=	O	=
<b>Grasp</b>	X	X	O	O	=	=	X	O	O	X	=	O	=	X	X	O	X	O	X	X
Obsolete	O	O	O	X	O	=	=	X	=	X	X	O	=	=	X	=	X	O	=	X

Overwhelm	X	=	O	O	X	=	X	X	=	=	=	=	=	=	O	=	X	=	=	X
Pant	X	X	=	X	O	X	X	X	X	X	O	X	O	X	O	X	X	=	X	
<b>Pepper</b>	X	O	X	X	X	X	X	X	X	X	X	=	X	=	=	X	X	X	X	
Perceive	X	X	X	X	=	=	=	O	=	X	=	=	=	=	O	=	O	=	=	=
<b>Pinprick</b>	X	X	X	X	=	=	=	X	=	X	X	=	=	O	X	=	=	=	=	X
Poise	O	X	X	X	O	X	O	X	X	X	=	=	O	X	X	=	O	X	X	X
Quizzically	X	X	X	=	=	X	X	X	X	X	=	=	=	=	O	=	X	=	=	=
Shiver	=	O	O	X	=	=	=	X	=	X	=	=	=	O	=	O	X	=	=	=
Shrug	X	X	X	X	X	=	X	X	=	X	=	=	=	=	X	=	X	=	=	X
Startle	X	O	X	X	=	X	X	O	=	X	X	=	=	=	X	O	X	=	=	X
Swirl	X	X	O	X	X	O	=	X	O	X	=	=	=	=	X	O	=	X	=	O
Tentatively	=	X	=	X	X	=	O	X	=	=	X	X	O	O	X	O	X	X	O	X
<b>Treat</b>	X	X	X	X	=	X	X	X	=	X	X	O	X	X	X	O	X	=	X	X
<b>Tug</b>	X	X	=	X	X	X	X	X	=	X	=	=	=	X	X	=	X	X	X	X
Unwieldy	=	X	X	O	X	=	O	X	O	X	=	=	O	=	X	=	X	O	=	O
Whirl	X	=	X	X	=	=	=	X	O	X	X	O	=	=	X	=	X	O	X	=
Wince	X	X	X	O	O	=	X	X	O	O	X	O	=	=	X	=	X	=	=	X

---

---

<b>Learned</b>	2	5	6	7	8	8	6	5	8	1	0	9	8	5	4	13	4	6	4	2
<b>Not-learned</b>	25	22	22	23	8	9	18	25	5	27	13	4	4	9	25	1	25	6	8	17
<b>Prior knowledge</b>	4	4	3	1	15	14	7	1	18	3	18	18	19	17	2	17	2	19	19	12
<b>Learned %</b>	6%	16%	19%	23%	26%	26%	19%	16%	26%	3%	0%	29%	26%	16%	13%	42%	13%	19%	13%	6%
<b>Not-learned %</b>	81%	71%	71%	74%	26%	29%	58%	81%	16%	87%	42%	13%	13%	29%	81%	3%	81%	19%	26%	55%
<b>Prior-knowledge %</b>	13%	13%	10%	3%	48%	45%	23%	3%	58%	10%	58%	58%	61%	55%	6%	55%	6%	61%	61%	39%

---

## Participants 61 ~ 80

Target words	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
Abruptly	=	X	O	=	=	X	X	X	X	X	O	X	X	O	X	X	X	X	X	X
Bask	X	X	X	=	O	O	=	=	X	O	O	X	=	X	=	X	X	O	X	X
Comprehend	=	O	X	X	=	X	=	X	X	X	=	X	=	=	X	=	=	=	X	X
Consciousness	=	X	X	=	=	X	=	X	X	X	X	O	O	X	O	=	O	=	X	X
Conveyance	=	O	O	X	=	X	=	O	X	X	X	X	=	X	X	X	=	=	X	X
<b>Crease</b>	X	X	O	X	X	O	O	X	X	X	X	X	X	X	O	X	O	X	X	X
Descent	X	X	X	X	=	X	=	O	=	X	X	X	X	X	O	X	X	=	X	X
<b>Drain</b>	X	O	X	O	=	X	O	X	X	X	O	X	=	=	X	=	=	X	X	O
Exhilarate	=	X	=	O	=	X	=	X	O	X	X	X	O	X	=	X	O	=	X	X
<b>Firmly</b>	=	X	X	=	=	=	=	O	X	X	X	=	O	X	X	=	O	O	O	X
Frigid	=	X	O	X	=	X	O	X	=	=	X	X	=	=	=	=	=	=	=	X
Glee	X	=	X	=	=	X	=	X	X	O	O	X	=	X	=	=	=	=	X	X
<b>Grasp</b>	X	X	=	X	=	X	O	X	X	X	X	=	X	=	=	X	X	X	=	X
Obsolete	=	X	X	X	=	O	=	X	X	=	X	X	=	X	X	=	O	O	=	X

Overwhelm	=	O	=	=	=	X	=	O	X	O	O	=	=	X	O	=	X	=	X	O
Pant	O	O	X	O	O	X	=	O	X	X	X	X	=	=	O	O	O	O	X	X
<b>Pepper</b>	O	X	O	O	=	=	X	X	X	X	=	O	X	X	X	X	X	X	X	=
Perceive	=	X	X	=	=	X	=	X	X	X	O	X	=	X	=	X	X	X	=	=
<b>Pinprick</b>	=	X	=	=	=	X	=	X	X	X	=	X	=	X	=	=	=	=	X	X
Poise	X	X	X	O	X	X	X	X	X	X	X	X	X	=	X	X	X	X	=	X
Quizzically	=	X	X	X	=	X	=	X	=	X	X	X	=	X	X	X	=	=	X	X
Shiver	O	=	=	=	=	X	=	X	=	X	O	X	=	X	=	O	=	=	X	O
Shrug	=	O	X	O	=	X	=	X	X	O	X	X	=	X	=	O	=	=	X	X
Startle	=	X	X	O	O	X	X	X	X	=	X	X	O	X	O	X	X	O	O	X
Swirl	X	X	=	O	X	X	X	O	X	X	X	X	=	X	X	X	O	O	X	X
Tentatively	=	X	X	X	=	=	O	X	=	=	X	X	X	X	X	X	X	=	X	X
<b>Treat</b>	=	=	=	=	=	X	=	X	X	X	X	X	=	X	X	X	X	X	=	X
<b>Tug</b>	=	=	=	X	=	X	=	O	O	O	=	X	=	X	X	X	X	X	X	X
Unwieldy	=	X	X	X	O	X	=	X	X	O	X	O	=	X	X	X	O	=	X	X
Whirl	=	=	O	X	=	=	X	X	X	=	X	O	=	X	=	X	X	X	X	X
Wince	X	X	O	X	=	O	O	O	X	X	O	X	O	X	X	X	O	O	O	O

---

---

<b>Learned</b>	3	6	7	8	4	4	6	8	2	6	8	4	5	1	6	3	9	7	3	4
<b>Not-learned</b>	9	20	16	13	3	23	6	22	24	20	19	24	7	24	15	19	13	10	22	25
<b>Prior knowledge</b>	19	5	8	10	24	4	19	1	5	5	4	3	19	6	10	9	9	14	6	2
<b>Learned %</b>	10%	19%	23%	26%	13%	13%	19%	26%	6%	19%	26%	13%	16%	3%	19%	10%	29%	23%	10%	13%
<b>Not-learned %</b>	29%	65%	52%	42%	10%	74%	19%	71%	77%	65%	61%	77%	23%	77%	48%	61%	42%	32%	71%	81%
<b>Prior-knowledge %</b>	61%	16%	26%	32%	77%	13%	61%	3%	16%	16%	13%	10%	61%	19%	32%	29%	29%	45%	19%	6%

---

## Participants 81 ~ 100

Target words	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Abruptly	X	X	=	X	=	X	=	X	X	X	X	O	=	=	=	X	X	X	X	X
Bask	X	O	=	=	O	=	=	=	O	X	X	=	O	O	X	O	X	X	X	X
Comprehend	=	X	=	X	=	=	=	O	X	X	X	X	O	=	=	=	=	X	X	=
Consciousness	X	X	=	X	=	X	X	X	X	X	O	=	O	=	=	X	=	=	O	X
Conveyance	=	O	O	=	X	X	X	X	X	X	O	X	X	X	X	X	X	X	X	X
<b>Crease</b>	X	X	X	=	X	X	O	O	X	X	X	X	O	=	=	=	X	=	=	=
Descent	X	X	=	X	X	O	X	X	X	O	O	X	=	=	X	O	=	=	=	X
<b>Drain</b>	X	X	=	X	X	X	=	X	X	X	X	X	X	X	X	X	=	=	X	X
Exhilarate	=	X	O	=	X	X	=	O	X	=	X	O	=	X	=	X	=	X	=	=
<b>Firmly</b>	X	X	O	X	X	X	O	X	=	X	X	X	X	=	=	X	O	=	=	X
Frigid	X	=	O	X	O	X	X	X	X	X	X	X	=	=	=	X	X	=	X	X
Glee	X	=	O	X	=	=	=	X	X	O	X	X	=	=	=	=	X	X	X	X
<b>Grasp</b>	O	X	X	=	X	=	=	X	X	X	X	O	=	=	X	=	=	=	X	X
Obsolete	=	=	O	=	X	X	X	O	X	O	X	X	=	X	=	X	X	X	X	X

Overwhelm	X	X	=	X	X	X	=	O	X	X	X	=	=	=	X	=	X	X	X	X
Pant	X	X	X	X	X	X	X	X	X	X	X	X	X	X	O	=	O	X	O	X
<b>Pepper</b>	X	X	X	X	X	X	X	X	X	X	=	X	X	X	X	X	X	X	X	X
Perceive	X	=	=	X	O	O	O	=	=	X	X	X	=	=	=	X	X	X	X	X
<b>Pinprick</b>	X	=	X	=	X	X	O	X	X	=	X	X	=	O	X	O	X	X	O	X
Poise	X	X	X	O	X	=	X	O	X	O	O	X	X	O	X	O	X	X	=	X
Quizzically	X	X	=	X	X	X	O	X	=	X	X	X	X	=	X	O	X	X	X	X
Shiver	=	X	O	X	=	X	X	X	X	=	X	X	O	=	=	X	X	X	X	X
Shrug	X	X	=	X	=	=	X	=	=	=	=	=	=	=	=	X	=	X	X	X
Startle	X	X	=	X	=	O	X	=	X	X	X	X	O	=	X	=	=	X	X	X
Swirl	X	O	O	X	X	X	X	X	O	X	X	X	=	X	X	X	X	X	X	X
Tentatively	X	X	=	X	=	X	X	X	X	X	X	O	X	X	X	X	O	X	X	X
<b>Treat</b>	=	=	=	X	X	X	X	X	X	O	X	=	X	X	X	X	X	X	=	X
<b>Tug</b>	X	X	X	=	X	O	=	X	X	X	X	O	X	X	O	=	X	X	O	X
Unwieldy	=	X	X	X	O	X	O	O	O	O	X	=	O	X	O	X	X	X	O	X
Whirl	X	X	=	X	=	O	X	X	=	X	X	X	O	X	X	=	X	X	X	X
Wince	=	X	X	X	O	X	X	X	=	X	X	X	O	=	O	=	X	X	X	X

---

---

<b>Learned</b>	1	3	8	1	5	5	6	7	3	6	4	5	9	3	4	5	3	0	5	0
<b>Not-learned</b>	22	22	9	22	17	20	16	20	22	21	25	20	10	12	15	16	20	24	20	28
<b>Prior knowledge</b>	8	6	14	8	9	6	9	4	6	4	2	6	12	16	12	10	8	7	6	3
<b>Learned %</b>	3%	10%	26%	3%	16%	16%	19%	23%	10%	19%	13%	16%	29%	10%	13%	16%	10%	0%	16%	0%
<b>Not-learned %</b>	71%	71%	29%	71%	55%	65%	52%	65%	71%	68%	81%	65%	32%	39%	48%	52%	65%	77%	65%	90%
<b>Prior-knowledge %</b>	26%	19%	45%	26%	29%	19%	29%	13%	19%	13%	6%	19%	39%	52%	39%	32%	26%	23%	19%	55%

---

\* POS: Part of Speech

\* Participants n = 100

\* Target words n = 31

\* Participant IDs were replaced with numbers 1 ~ 100

Item level summary ( n = 100)

Target words	Learned	Not-learned	Prior knowledge
Abruptly	12	70	18
Bask	32	46	22
Comprehend	8	43	49
Consciousness	12	47	41
Conveyance	18	54	28
<b>Crease</b>	29	59	12
Descent	19	58	23
<b>Drain</b>	15	51	34
Exhilarate	21	48	31
<b>Firmly</b>	27	44	29
Frigid	20	42	38
Glee	17	48	35
<b>Grasp</b>	16	45	39
Obsolete	21	50	29
Overwhelm	12	40	48

Pant	27	60	13
<b>Pepper</b>	7	76	17
Perceive	13	46	41
<b>Pinprick</b>	11	49	40
Poise	17	70	13
Quizzically	7	50	43
Shiver	16	40	44
Shrug	17	42	41
Startle	19	58	23
Swirl	20	64	16
Tentatively	18	56	26
<b>Treat</b>	10	63	27
<b>Tug</b>	12	58	30
Unwieldy	23	46	31
Whirl	12	55	33
Wince	26	52	22

---